Year: 2019

Cosmological information contents on the light-cone

Yoo, Jaiyul ; Mitsou, Ermis ; Grimm, Nastassia ; Durrer, Ruth ; Refregier, Alexandre

Abstract: We develop a theoretical framework to describe the cosmological observables on the past light cone such as the luminosity distance, weak lensing, galaxy clustering, and the cosmic microwave background anisotropies. We consider that all the cosmological observables include not only the background quantity, but also the perturbation quantity, and they are subject to cosmic variance, which sets the fundamental limits on the cosmological information that can be derived from such observables, even in an idealized survey with an infinite number of observations. To quantify the maximum cosmological information content, we apply the Fisher information matrix formalism and spherical harmonic analysis to cosmological observations, in which the angular and the radial positions of the observables on the light cone carry different information. We discuss the maximum cosmological information that can be derived from five different observables: (1) type Ia supernovae, (2) cosmic microwave background anisotropies, (3) weak gravitational lensing, (4) local baryon density, and (5) galaxy clustering. We compare our results with the cosmic variance obtained in the standard approaches, which treat the light cone volume as a cubic box of simultaneity. We discuss implications of our formalism and ways to overcome the fundamental limit.

DOI: https://doi.org/10.1088/1475-7516/2019/12/015

Originally published at:
Yoo, Jaiyul; Mitsou, Ermis; Grimm, Nastassia; Durrer, Ruth; Refregier, Alexandre (2019). Cosmological information contents on the light-cone. Journal of Cosmology and Astroparticle Physics, 2019(12):015.
DOI: https://doi.org/10.1088/1475-7516/2019/12/015

Cosmological Information Contents on the Light-Cone

Jaiyul Yoo,[a,b] Ermis Mitsou,[a] Nastassia Grimm,[a] Ruth Durrer[c] and Alexandre Refregier[d]

[a] *Center for Theoretical Astrophysics and Cosmology, Institute for Computational Science University of Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland*

[b]*Physics Institute, University of Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland*

[c]*Département de Physique Théorique & Center for Astroparticle Physics, Université de Genève Quai E. Ansermet 24, CH-1211 Genève 4, Switzerland*

[d]*Institute for Particle Physics and Astrophysics, Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8057 Zürich, Switzerland*

May 22, 2019

Abstract

We develop a theoretical framework to describe the cosmological observables on the past light cone such as the luminosity distance, weak lensing, galaxy clustering, and the cosmic microwave background anisotropies. We consider that all the cosmological observables include not only the background quantity, but also the perturbation quantity, and they are subject to cosmic variance, which sets the fundamental limits on the cosmological information that can be derived from such observables, even in an idealized survey with an infinite number of observations. To quantify the maximum cosmological information content, we apply the Fisher information matrix formalism and spherical harmonic analysis to cosmological observations, in which the angular and the radial positions of the observables on the light cone carry different information. We discuss the maximum cosmological information that can be derived from five different observables: (1) type Ia supernovae, (2) cosmic microwave background anisotropies, (3) weak gravitational lensing, (4) local baryon density, and (5) galaxy clustering. We compare our results with the cosmic variance obtained in the standard approaches, which treat the light cone volume as a cubic box of simultaneity. We discuss implications of our formalism and ways to overcome the fundamental limit.

[1]jyoo@physik.uzh.ch, ermitsou@physik.uzh.ch, ngrimm@physik.uzh.ch, ruth.durrer@unige.ch, alexandre.refregier@phys.ethz.ch

Contents

1 Introduction

The perturbations in cosmological large-scale structure are generated by the quantum fluctuations in the early Universe. On large scales, they are almost fully characterized by a Gaussian distribution and its power spectrum. Due to the random nature of the perturbations, many independent modes need to be sampled to obtain a robust estimate of the underlying power spectrum. Given a survey volume, however, there exists a finite number of independent modes available for large-scale observations, and these observations are subject to the large-scale fluctuations, known as sample variance. When the survey volume is limited by our observable Universe, sample variance is called cosmic variance, and it sets a limit to the fundamental cosmological information contents available to us (see, e.g., [1–4]). In the past, when the large-scale observations were limited to a small sky coverage and redshift range, the survey volume was well approximated as a cubic box of simultaneity, far away from the observer, so that the traditional 3D Fourier analysis provided a useful way to estimate the cosmological information contents in the survey. As we assume statistical homogeneity, each Fourier mode in a rectangular volume is independent, and its sample variance is set by its own power spectrum [5, 6]. However, experimental and observational techniques have developed rapidly in recent years and the angular coverage and redshift depth of large-scale surveys has become wider and deeper, so that this simple approximation of estimating the cosmic variance needs to be revisited.

Cosmological observables are mapped on the observer sky, always in terms of angular positions on the sky and redshifts. In particular, the radial position obtained by the *observed* redshift carries different information — the observers locate the cosmological observables along the past light cone with the redshift, while the angular position of the cosmological observables spans the two-dimensional sphere of constant redshift seen by the observer. Therefore, it is evident that the traditional Fourier analysis in a rectangular box is fundamentally incompatible with how the cosmological observables are mapped in the observer sky. This inadequacy is maximally manifest in the analysis of the cosmic microwave background (CMB) anisotropies, where the sky coverage is (almost) a full sphere and the angular positions of the CMB temperature and polarization anisotropies are decomposed in terms of spherical harmonics with discrete angular momentum, rather than with continuous Fourier modes. Moreover, significant progress has been made in the large-scale galaxy surveys to map the three-dimensional distribution of the matter density, greatly improving upon the first-generation surveys with the volume $\ll 0.1$ Gpc^3. In particular, the upcoming stage-IV surveys such as the Dark Energy Spectroscopic Instrument [7] and Large Synoptic Survey Telescope [8] and two space missions, Euclid [9] and the Wide Field Infrared Survey Telescope [10], will measure millions of galaxies with great precision by observing together almost a half of the entire sky over a large range of redshift. In this era of precision cosmology, the traditional Fourier analysis is increasingly inaccurate and becomes the source of systematic errors.

On the observed light cone we assume statistical isotropy so that the spherical harmonic modes are independent. However, the radial direction on the lightcone is mixed with time evolution which breaks translation symmetry. Therefore large radial modes are not statistically independent and we expect cross correlation between different redshifts. Depending on the observable considered, these can be very relevant and contain important cosmological information.

Here, we develop a theoretical framework to generally describe cosmological observables on the light cone and use the Fisher information matrix to quantify the maximum cosmological information contents available from observations of such cosmological observables. For two-dimensional angular observables such as CMB anisotropies, the standard formalism based on spherical harmonics is as accurate as our new formalism, except for one subtlety that background quantities cannot be measured by observations from a single light cone due to the perturbation of the monopole. This subtlety is often ignored, leading to an interesting bias and information loss, as we detail in section 5.2. For three-dimensional observables, the spherical Fourier analysis based on the radial and angular eigenfunctions of the Helmholtz equation is well developed [11–13], and it has been applied to the observational data analysis [14, 15] and to theoretical predictions [16–21]. Complementing the Fourier analysis to the spherical harmonics decomposition, the spherical Fourier analysis provides the most natural way to analyze the cosmological observables in the observer sky. However, the difficulty lies in computing the inverse spherical power spectrum, needed to quantify cosmic variance. As opposed to the angular analysis or the traditional Fourier analysis, where the inverse can be trivially obtained, the inverse spherical power spectrum has not been derived in the spherical Fourier analysis. Here we derive the inverse spherical power spectrum and use it to compute the maximum cosmological information contents under the assumption of Gaussianity. A similar idea was pursued [22] to compute the cosmological information contents as a function of the cosmic time, though, under the simplifying assumptions using the traditional Fourier analysis.

The organization of the paper is as follows: We first develop a unified theoretical framework for modeling the three-dimensional cosmological observables, the angular observables, and the projected observables in section 2. In section 3, we derive the likelihood of the cosmological observables on the light cone under the assumption of Gaussianity. We then use the Fisher information technique in section 4 to quantify the maximum cosmological information contents available from the cosmological observables. In section 5, we apply our Fisher matrix calculations to the luminosity distance, the CMB anisotropies, 3D weak lensing, the cosmic baryon density measurements, and the galaxy power spectrum. We discuss the implications of our new formulation of the cosmic variance on the light cone in section 6. In Appendix A, we present an alternative to the usual spherical Fourier analysis, in which the radial position is decomposed directly in terms of the observed redshift, avoiding the need to rely on a fiducial cosmology. In Appendix B, we present the relation of the spherical power spectrum on the light cone to the usual power spectrum on a hypersurface of simultaneity.

2 Cosmological Observables on the Light Cone

Here we present the theoretical descriptions of cosmological observables on the light cone. We start with the most important observable, namely, number counts of luminous objects, which we collectively call galaxies. Next, we provide theoretical descriptions of other cosmological observables derived from observations of galaxies such as the luminosity distance, the weak lensing shear, and so on. Drawing on these theoretical descriptions of cosmological observables, we consider two additional cases, in which the observations are limited to a single redshift bin and the observations are projected along the line-of-sight direction.

The main point of our formalism is to account for the fact that the cosmological observables are obtained only on the light cone volume, rather than on a hypersurface of simultaneity. Compared to the standard descriptions, this consideration significantly changes our theoretical descriptions of the cosmological observables in this section and more dramatically the cosmological information contents in section 4.

2.1 Galaxy clustering and number counts

Luminous objects such as galaxies are easy to observe up to very high redshift, providing a great opportunity for us to probe the Universe. In particular, their number density is the primary observable and its two-point correlation (or the power spectrum) has been widely used to test our theoretical models and to understand the Universe (see, e.g., [23] for review). Galaxy clustering contains a wealth of cosmological information. Its intrinsic correlation encodes the underlying matter distribution, and the volume effects involve the redshift space distortion and gravitational lensing (see, e.g., [24–26]), in addition to subtle relativistic effects.

In observations, we find a luminous object given the conditions for its color and morphology and the thresholds for its brightness in the simplest case, and the object that satisfies the conditions is identified as a galaxy within a redshift bin $(z, z+dz)$ and a solid angle $d\Omega = \sin\theta d\theta d\phi$. Its position is then recorded in terms of the redshift z and angular position $\hat{\mathbf{n}} = (\theta, \phi)$, which we represent by $\mathbf{x}_i$

for i-th galaxy. For its theoretical description, the observers often assume a cosmological model to convert it in the observer frame:

$$\mathbf{x} := \bar{r}(z)\hat{\mathbf{n}} = \bar{r}(z)(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \ . \tag{2.1}$$

Throughout the paper, we often use $\mathbf{x} = (z, \hat{\mathbf{n}})$ for notational simplicity. The total number N of the observed galaxies,

$$N = \sum_{i=1}^{N} dN(\mathbf{x}_i) \ , \qquad\qquad dN \in \{0, 1\} \ , \tag{2.2}$$

is simply the sum of the number dN of galaxies in a small volume $d\bar{V}$ centered at $\mathbf{x}_i$. The number count $dN(\mathbf{x}_i)$ is then further related to the (observed) galaxy number density $n(\mathbf{x}_i)$ as

$$dN(\mathbf{x}_i) = \frac{dN}{dzd\Omega}(\mathbf{x}_i)dzd\Omega := n(\mathbf{x}_i)d\bar{V}(\mathbf{x}_i) \ , \tag{2.3}$$

where the (observed) background volume element,

$$d\bar{V}(z_i) = \frac{\bar{r}^2(z_i)}{H(z_i)(1+z_i)^3}\, dz\, d\Omega \ , \tag{2.4}$$

is the physical volume, corresponding to the observed redshift bin dz and the solid angle $d\Omega$ in an assumed homogeneous background universe. Given a set of cosmological parameters, the observers can convert the redshift and angle to physical distances by using the Hubble parameter $H(z)$ and the angular diameter distance

$$\bar{r}(z) = \int_0^z \frac{dz'}{H(z')} \ , \tag{2.5}$$

where we assume a flat universe ($K = 0$).

To model the observations, we develop theoretical predictions that describe such observations of galaxies at any (continuous) point $\mathbf{x}$ within the survey volume (instead of discrete observation points $\mathbf{x}_i$), while keeping the total number N of the observed galaxies. The observed galaxy number density is now modeled as a continuous number density field

$$n(\mathbf{x}) := \bar{n}(z)\left[1 + \delta_g(\mathbf{x})\right] \ , \tag{2.6}$$

where we split the number density into the number density $\bar{n}(z)$ in the background and its perturbation δ_g around the background. Furthermore, since the inhomogeneities in the Universe affect our cosmological observables, the observed galaxy number density $n(\mathbf{x})$ and the volume element $d\bar{V}$ are different from the physical galaxy number density n_p and the physical volume dV_p in the inhomogeneous Universe that corresponds to the observed position $\mathbf{x} = (z, \hat{\mathbf{n}})$. Indeed, the observed number count is related as

$$dN(\mathbf{x}) = n_p(\mathbf{x})dV_p(\mathbf{x}) := n(\mathbf{x})d\bar{V}(\mathbf{x}) \ , \tag{2.7}$$

and these physical quantities can be further split as

$$n_p(\mathbf{x}) := \bar{n}(z)\left[1 + \delta_s(\mathbf{x})\right] \ , \qquad\qquad dV_p(\mathbf{x}) := d\bar{V}(\mathbf{x})\left[1 + \delta V(\mathbf{x})\right] \ , \tag{2.8}$$

where δ_s is the intrinsic fluctuation of the galaxy number density and δV is the (dimensionless) fluctuation of the physical volume element. From Eq. (2.7), the observed galaxy (number density) fluctuation is then derived as

$$\delta_g(\mathbf{x}) = \delta_s(\mathbf{x}) + \delta V(\mathbf{x}) + \delta_s(\mathbf{x})\delta V(\mathbf{x}) \ , \tag{2.9}$$

and proper relativistic computations of the gauge-invariant expression δ_g and its correlations have been the focus of recent research (see, e.g., [27–35] for details), as it involves subtle gauge issues and they can be used to extract extra cosmological information. For our present purposes, we do not need to know their detailed expressions, but it suffices to note two things: (1) Though the above expressions are exact, the individual quantities are perturbations, such that the quadratic terms can be dropped for the linear-order calculations. (2) The individual terms are gauge-invariant and expressed at the observed position $\mathbf{x}$.

Before we proceed, we define a few more perturbation quantities associated with the observed number counts. First, we define the background redshift distribution

$$\frac{d\bar{N}}{dzd\Omega}(z) := \frac{\bar{n}(z)\bar{r}^2(z)}{H(z)(1+z)^3} , \tag{2.10}$$

which is the number of galaxies we would measure per redshift bin dz and solid angle $d\Omega$ in a homogeneous universe. The observed number count is then modeled as a continuous field

$$dN(\mathbf{x}) = \bar{n}(z)d\bar{V}(\mathbf{x})\,[1+\delta_g(\mathbf{x})] = \frac{d\bar{N}}{dzd\Omega}(z)dzd\Omega\,[1+\delta_g(\mathbf{x})] \ , \tag{2.11}$$

and the total number of the observed galaxies is simply

$$\begin{aligned} N(z_{\max}) \ &= \int dN = \int_0^{z_{\max}} dz\, \frac{d\bar{N}}{dzd\Omega} \int d\Omega\,(1+\delta_g) \\ &:= \bar{N}(z_{\max}) + 4\pi \int_0^{z_{\max}} dz\, \frac{d\bar{N}}{dzd\Omega}\ \langle\delta_g\rangle_\Omega := \bar{N}\,(1+\delta N) \ , \end{aligned} \tag{2.12}$$

where we defined the background total number $\bar{N}$ and its dimensionless fluctuation δN (both of them are independent of direction)

$$\bar{N}(z_{\max}) := 4\pi \int_0^{z_{\max}} dz\, \frac{d\bar{N}}{dzd\Omega} , \qquad\qquad \delta N(z_{\max}) := 4\pi \int_0^{z_{\max}} dz\, \frac{1}{\bar{N}}\frac{d\bar{N}}{dzd\Omega}\ \langle\delta_g\rangle_\Omega \ , \tag{2.13}$$

and the angle-averaged galaxy fluctuation (or the monopole)

$$\langle\delta_g\rangle_\Omega(z) := \int \frac{d\Omega}{4\pi}\,\delta_g(\mathbf{x}) = \int \frac{d\Omega}{4\pi}\,[\delta_s(\mathbf{x}) + \delta V(\mathbf{x}) + \delta_s(\mathbf{x})\delta V(\mathbf{x})] \ . \tag{2.14}$$

Note that only the total number N of the observed galaxies is a physical number and the split of N into $\bar{N}$ and δN is purely theoretical for later convenience. Here we assume a full sky coverage for simplicity.

2.2 Other cosmological observables

Beyond the primary observable or the galaxy number counts, other cosmological information can be extracted from the observations of luminous objects such as the luminosity distance from type Ia supernovae, the lensing shear from the shape of galaxies, and so on. These cosmological observables can be used to compute their two-point correlation (or higher statistics) in the same way galaxy clustering is measured, and they contain equally important cosmological information, compared to measurements of galaxy clustering.

In addition to the galaxy number counts, we consider an observable quantity D from galaxies such as the luminosity distance and construct the observed data set $\mathcal{D}_i^{\rm obs}$ for i-th galaxy at $\mathbf{x}_i$:

$$\boldsymbol{\mathcal{D}}^{\rm obs} = \{D(\mathbf{x}_1),\ D(\mathbf{x}_2),\ \cdots,\ D(\mathbf{x}_N)\} , \tag{2.15}$$

where we used the boldface letter to indicate the observed data set is a vector. As we are often interested in the background quantity $\bar{D}(z)$ at a given redshift z, we may average the observed data set to derive an estimate of $\bar{D}(z)$, if all the measurements are in the same redshift bin:

$$\langle \mathcal{D} \rangle^{\mathrm{obs}}(z) := \frac{1}{N}\sum_{i=1}^{N} \mathcal{D}_i^{\mathrm{obs}} \,, \qquad\qquad z_i \in (z, z+dz) \,, \tag{2.16}$$

where we define the notation for the average $\langle \cdots \rangle$ of observed quantities. In the same spirit, we model these observations and develop theoretical predictions that provide the observable $D(\mathbf{x})$ from galaxies at any (continuous) point $\mathbf{x}$ (instead of discrete observation points $\mathbf{x}_i$), and the observable quantity D is then approximated as a continuous field

$$D(\mathbf{x}) := \bar{D}(z)\left[1 + \delta D(\mathbf{x})\right] \,, \tag{2.17}$$

where we again split the observable quantity D into a background $\bar{D}$ and the dimensionless perturbation δD around the background. The perturbation δD is a diffeomorphism invariant scalar, and it is gauge-invariant at the linear order [36]. For example, the perturbation δD for the luminosity distance as an observable quantity D has been computed (see, e.g., [37–42]), and it involves the Doppler effects and subtle relativistic effects associated with the light propagation. Here, we leave it general as the detailed expressions are less important for our current discussion.

Note, however, that the continuous observable field $D(\mathbf{x})$ alone cannot be a complete description of our observed data set $\mathcal{D}_i^{\mathrm{obs}}$ in the continuous limit, as we measure the observable D only through observations of galaxies at $\mathbf{x}_i$. The observed data set $\mathcal{D}_i^{\mathrm{obs}}$ in Eq. (2.15) is not obtained by uniform sampling of D over the survey volume, but by sampling biased objects such as galaxies. For example, considering the luminosity distance observations from supernovae, we have fewer measurements of the luminosity distance (or none) in an underdense region, where there are fewer galaxies and supernovae, despite the fact that the luminosity distance to this underdense region is non-zero. The observed data set $\mathcal{D}_i^{\mathrm{obs}}$ in Eq. (2.15) is indeed a set of observables weighted by the galaxy number counts, such that our theoretical description is

$$\mathcal{D}(\mathbf{x}) = D(\mathbf{x})\frac{dN(\mathbf{x})}{N} \,, \tag{2.18}$$

where $dN(\mathbf{x})$ becomes zero or one in the observed data set and we simply divided by the (constant) total number N of galaxies for later convenience. For instance, the angle average of $D(\mathbf{x})$ does not reproduce the observed angle average in Eq. (2.16), but the angle average of $\mathcal{D}(\mathbf{x})$ does, as in Eq. (2.23). Finally, we split our model for the observed data set

$$\mathcal{D}(\mathbf{x}) := \bar{\mathcal{D}}(z)\left[1 + \delta\mathcal{D}(\mathbf{x})\right] \,, \tag{2.19}$$

into the background and the perturbation parts

$$\bar{\mathcal{D}}(z) = \frac{\bar{D}(z)}{\bar{N}}\frac{d\bar{N}}{dzd\Omega}(z)dzd\Omega = \frac{\bar{D}(z)\bar{n}(z)\bar{r}^2(z)}{\bar{N}H(z)(1+z)^3}dzd\Omega =: \hat{\bar{\mathcal{D}}}(z)dzd\Omega \,, \tag{2.20}$$

$$\delta\mathcal{D}(\mathbf{x}) = \frac{\left[1+\delta D(\mathbf{x})\right]\left[1+\delta_g(\mathbf{x})\right]}{1+\delta N} - 1 = \frac{\delta D(\mathbf{x}) + \delta_g(\mathbf{x}) - \delta N + \delta D(\mathbf{x})\delta_g(\mathbf{x})}{1+\delta N} \,. \tag{2.21}$$

We include the observational bin sizes for the redshift dz and the solid angle $d\Omega$ in the background $\bar{\mathcal{D}}(z)$, as they are set by observers. The dimensionless fluctuation $\delta\mathcal{D}$ in our observed data set is driven, not only by the fluctuation δD in the observable D, but also by the fluctuation δ_g in the observed galaxy number density. In case there is no galaxy ($\delta_g = -1$) at a given position and hence no observation at all, the perturbation part is $\delta\mathcal{D} = -1$, implying no observed data $\mathcal{D} = 0$, consistent with our arguments. As stated in the example of the luminosity distance observable, the continuous field $\mathcal{D}(\mathbf{x})$ as our theoretical modeling of the observed data set should not be considered as the luminosity distance itself at a given position, but as a useful description of the observed data set in the limit $N \to \infty$, accounting for the bias due to the host galaxy clustering.

When our cosmological observable is simply the galaxy number density n, we can directly use Eq. (2.19) for galaxy clustering by replacing $\bar{D} = 1$ and setting $\delta D = 0$ in Eqs. (2.20) and (2.21), where no further weight by the galaxy number counts is needed. We summarize our notation in Table 1.

2.3 Observables in a single redshift bin

It is often the case that the observed data set is confined to a single redshift bin $z_i \in [z, z + \Delta z]$, as one considers a sub-sample at the single redshift bin out of the full observations, spanning a wide range of redshift. In this case, our theoretical descriptions of the data can be obtained by replacing the infinitesimal redshift bin size dz with the finite (constant) width Δz as

$$\bar{N} = 4\pi\Delta z \frac{d\bar{N}}{dz d\Omega}\,, \qquad \bar{\mathcal{D}}(z) = \bar{D}(z)\frac{d\Omega}{4\pi} =: \hat{\bar{\mathcal{D}}}(z) d\Omega\,, \qquad \mathcal{D}(\mathbf{x}) = \bar{D}(z)\frac{d\Omega}{4\pi}\left[1 + \delta\mathcal{D}(\mathbf{x})\right]\,, \tag{2.22}$$

where Δz drops out in $\bar{\mathcal{D}}$ and $\mathcal{D}$ due to Δz in $\bar{N}$. The angular average

$$\langle \mathcal{D} \rangle_\Omega (z) := \bar{D}(z) \int \frac{d\Omega}{4\pi}\left[1 + \delta\mathcal{D}(\mathbf{x})\right]\,, \tag{2.23}$$

can be compared to the observed average $\langle \mathcal{D} \rangle^{\text{obs}}$ in Eq. (2.16). Note that the total number N of observed galaxies as the denominator in Eq. (2.16) is already accounted for in our theoretical description.

At linear order in perturbations, the angle average of $\delta\mathcal{D}(\mathbf{x})$ is often equated with an ensemble average and is ignored as its perturbation vanishes at the linear order. However, the angle average

$$\delta\mathcal{D}_0(z) := \int \frac{d\Omega}{4\pi}\, \delta\mathcal{D}(z, \hat{\mathbf{n}}) \neq 0 \tag{2.24}$$

is non-vanishing even at the linear order in perturbations, unless it is further averaged over all the possible observer positions in the Universe, or other light cones [43], assuming the Ergodic hypothesis. Indeed, the angle average is referred to as the monopole perturbation from the general multipole expansion

$$\delta\mathcal{D}(\mathbf{x}) = \sum_{lm} a_{lm}(z) Y_{lm}(\hat{\mathbf{n}})\,, \qquad \delta\mathcal{D}_0(z) = \frac{a_{00}}{\sqrt{4\pi}}\,. \tag{2.25}$$

In the standard redshift-space distortion power spectrum analysis (see, e.g., [25, 44, 45]), the observed galaxy fluctuation $\delta_z(\mathbf{x})$ is fully decomposed into the monopole, the quadrupole, and the hexadecapole, none of which are zero.[2]

The cosmic microwave background (CMB) anisotropies also correspond to the case of observables in a single redshift bin, in which the observed redshift is zero (no measurements in other redshifts, in practice):

$$\bar{D}(z) := \bar{T}\,, \qquad \delta\mathcal{D}(\mathbf{x}) \equiv \delta D(\mathbf{x}) \equiv \Theta(\hat{\mathbf{n}}) := \frac{\delta T(\hat{\mathbf{n}})}{\bar{T}}\,. \tag{2.26}$$

[2]Note, however, that the multipoles in the redshift-space distortion are to be understood as those of $\hat{\mathbf{k}}$ with respect to a fixed observer direction $\hat{\mathbf{n}}$, while we decompose our variables as functions of $\hat{\mathbf{n}}$ into their multipoles, hence these multipoles mean something very different.

CMB anisotropies are an unbiased observable ($\delta_g \equiv 0$), i.e no weight δ_g from the galaxy number counts is involved in the description. However, there exists a subtlety associated with the background CMB temperature $\bar{T}$ at $z = 0$, which is an input cosmological parameter of the model that cannot be exactly determined by observations. The observed CMB temperature $\langle T\rangle^{\rm obs}$ in observations [46] is indeed the angular average in Eqs. (2.23) and (2.16) of the temperature measurements on the sky, and it is different from the background temperature $\bar{T}$ again due to the monopole contribution Θ_0 of the perturbation. The same difference exists in the luminosity distance, for example, between the background luminosity distance $\bar{D}_L(z)$ and the angle average $\langle\mathcal{D}_L(z)\rangle_\Omega$. The former is a mathematical function of a homogeneous and isotropic universe, and the latter involves the perturbations from various effects. This difference arises as we have a single light cone in the Universe [43]. Note also that the larger the sphere over which we take an angle average the smaller its 'cosmic variance'. Hence cosmic variance is most substantial at small redshift.

In the CMB literature, the observed CMB temperature $\langle T\rangle^{\rm obs}$ is always used, so that it is more convenient to re-arrange the expression for the CMB temperature on the sky:

$$T(\hat{\mathbf{n}}) = \bar{T}\left[1+\Theta(\hat{\mathbf{n}})\right] = \bar{T}\left(1+\Theta_0\right) \times \frac{1+\Theta(\hat{\mathbf{n}})}{1+\Theta_0} \equiv \langle T\rangle^{\rm obs}\left[1+\Theta^{\rm obs}(\hat{\mathbf{n}})\right] , \tag{2.27}$$

which defines the observed temperature and its anisotropies

$$\langle T\rangle^{\rm obs} := \frac{1}{N}\sum_{i=1}^{N} T(\hat{\mathbf{n}}_i) = \int \frac{d\Omega}{4\pi}\, T(\hat{\mathbf{n}}) \equiv \bar{T}(1+\Theta_0) , \qquad 1+\Theta^{\rm obs}(\hat{\mathbf{n}}) := \frac{1+\Theta(\hat{\mathbf{n}})}{1+\Theta_0} . \tag{2.28}$$

Note that since the monopole Θ_0 is already absorbed in $\langle T\rangle^{\rm obs}$, there is no monopole $\Theta_0^{\rm obs} \equiv 0$ in observations. In section 5.2 we further discuss subtleties in $\Theta(\hat{\mathbf{n}})$ associated with gauge choice and observer frame. Table 1 summarizes our theoretical description of the observed data set in a single redshift bin.

2.4 Angular (projected) observables

In weak lensing observations, the cosmological observables D are primarily composed of the lensing convergence κ, shear γ_1, γ_2, and rotation ω. The proper relativistic description of the lensing observables has been developed recently [47, 48], demonstrating the existence of additional relativistic effects missing in the standard weak lensing formalism and resolving the subtle issues associated with gauge choice and physical rotation (see also [49–52]). Here, we are not concerned with the details of $\delta\mathcal{D}$, but we note that the lensing observables vanish in the background due to the symmetry. Therefore, the continuous field of the observable quantity D in Eq. (2.17) is in this case modeled as

$$D(\mathbf{x}) := \delta D(\mathbf{x}) , \tag{2.29}$$

without a background part, and our theoretical description of the observed data set is then described as

$$\mathcal{D}(\mathbf{x}) = \delta D(\mathbf{x})\frac{dN(\mathbf{x})}{N} =: \bar{\mathcal{D}}(z)\delta\mathcal{D}(\mathbf{x}) , \tag{2.30}$$

where we defined the "background"

$$\bar{\mathcal{D}}(z) = \frac{1}{\bar{N}}\frac{d\bar{N}}{dzd\Omega}(z)dzd\Omega = \frac{\bar{n}(z)\bar{r}^2(z)}{\bar{N}H(z)(1+z)^3}dzd\Omega , \tag{2.31}$$

Table 1. Theoretical descriptions of the cosmological observables

observable $\mathcal{D} = \bar{\mathcal{D}}(z)(1+\delta\mathcal{D})$	$\bar{\mathcal{D}}(z)$	$\delta\mathcal{D}(\mathbf{x})$	Equations
three-dimensional observables	$\frac{\bar{D}(z)\bar{n}(z)\bar{r}^2(z)}{\bar{N}H(z)(1+z)^3}dzd\Omega$	$\frac{\delta D(\mathbf{x})+\delta_g(\mathbf{x})-\delta N+\delta D(\mathbf{x})\delta_g(\mathbf{x})}{1+\delta N}$	(2.20), (2.21)
galaxy clustering ($\bar{D}=1$, $\delta D=0$)	$\frac{\bar{n}(z)\bar{r}^2(z)}{H(z)(1+z)^3}dzd\Omega$	$\delta_g(\mathbf{x})$	(2.20), (2.21)
single redshift bin ($dz=\Delta z$)	$\frac{\bar{D}(z)}{4\pi}d\Omega$	$\frac{\delta D(\mathbf{x})+\delta_g(\mathbf{x})-\delta N+\delta D(\mathbf{x})\delta_g(\mathbf{x})}{1+\delta N}$	(2.22), (2.21)
CMB $T(\hat{\mathbf{n}})=\bar{T}(1+\Theta)$	$\bar{T}\frac{d\Omega}{4\pi}$	$\Theta(\hat{\mathbf{n}})$	(2.26)
lensing observable $\mathcal{D}=\bar{\mathcal{D}}(z)\delta\mathcal{D}$	$\frac{\bar{n}(z)\bar{r}^2(z)}{\bar{N}H(z)(1+z)^3}dzd\Omega$	$\frac{\delta D(\mathbf{x})[1+\delta_g(\mathbf{x})]}{1+\delta N}$	(2.30)−(2.32)

and the perturbation $\delta\mathcal{D}$ (but instead of $1+\delta\mathcal{D}$)

$$\delta\mathcal{D}(\mathbf{x}) = \frac{\delta D(\mathbf{x})\,[1+\delta_g(\mathbf{x})]}{1+\delta N}\,. \tag{2.32}$$

The perturbation part $\delta\mathcal{D} = \delta D + \mathcal{O}(2)$ is devoid of any contribution of the galaxy number density fluctuation δ_g at the linear order, and $\bar{\mathcal{D}}(z)$ is nothing but the (normalized) redshift distribution of the observed galaxy number density in the background.

Since these observables like the lensing observables are often dominated by measurement errors such as shape noise or redshift measurement errors, they are summed over the line-of-sight direction to enhance the signal-to-noise ratio. This procedure is simply accommodated in our theoretical description as the projected observed data set

$$\mathcal{D}^{\rm pro}(\hat{\mathbf{n}}) := \int_z \mathcal{D}(\mathbf{x}) = \int dz\left(\frac{1}{\bar{N}}\frac{d\bar{N}}{dzd\Omega}\right)d\Omega\;\delta\mathcal{D}(\mathbf{x})\,, \tag{2.33}$$

where we use the superscript to represent the quantities that are projected along the line-of-sight direction. Similarly, the projected galaxy number density is often used to probe galaxy clustering, when the redshift measurements are obtained by noisier photo-z measurements or for radio galaxies where redshift information is very uncertain. The projected number $dN^{\rm pro}(\hat{\mathbf{n}})$ of the observed galaxies in a solid angle $d\Omega$ is simply related to the number of the observed galaxies in a volume dV, but projected as

$$dN^{\rm pro}(\hat{\mathbf{n}}) = \int_z dN(\mathbf{x}) = \int_z n(\mathbf{x})dV = d\Omega\int dz\frac{dN}{dzd\Omega} =: n^{\rm 2D}(\hat{\mathbf{n}})d\Omega\,, \tag{2.34}$$

where we introduce the angular galaxy number density $n^{\rm 2D}$ and note that the angular galaxy number density is dimensionless. Therefore, our theoretical description of the angular galaxy number density is

$$n^{\rm 2D}(\hat{\mathbf{n}}) = \int dz\frac{d\bar{N}}{dzd\Omega}\,[1+\delta_g(\mathbf{x})]\,. \tag{2.35}$$

In case that our cosmological observable is simply the projected galaxy number density $n^{\rm 2D}$, we can directly use Eq. (2.33) for angular galaxy clustering by and setting $\delta D = 1$ in Eq. (2.32). Table 1 summarizes our theoretical description of the angular observables.

3 Likelihood of Cosmological Observables on the Light Cone

Given the redshift z and the angular position $\hat{\mathbf{n}}$ of galaxies and the cosmological observables $\mathcal{D}$ derived from the other properties of galaxies such as the luminosity distance, lensing shear, and so on, we construct our favorite statistics to test the underlying cosmological models, and the tests of our cosmological models against the observations are performed through the likelihood analysis: A given set of cosmological parameters is used to predict the likelihood of the measurements, and combined with the priors of the cosmological parameters, the posterior is computed for each cosmological parameter set, and this procedure is repeated until the best parameter set that maximizes the posterior is found (see, e.g., [53, 54]).

This likelihood analysis is generic and standard in literature. Here we reformulate the likelihood analysis, taking into consideration that the cosmological observables are obtained on the light cone. This can be contrasted to the standard method in literature, in which the analysis is performed as though the survey volume would be a hypersurface of simultaneity. As long as the survey volume is small, the standard method is a good approximation to the real observations, but the systematic errors grow as redshift depth and the sky coverage increase. Here we develop the likelihood analysis on the light cone without such limitations. In particular, we will compute the *maximum information* contained on the light cone.

3.1 Gaussian probability distribution

Though our formalism is generally applicable to the likelihood analysis on the light cone, we make a series of assumptions to simplify our analytic calculations: From now on, we will exclusively deal with the linear perturbations, ignoring any higher-order perturbations, and we further assume that the linear perturbations have vanishing statistical mean and they are Gaussian-distributed on average. Under these assumptions, the two-point correlation will contain all the information, and any connected N-point correlations with $N > 2$ vanish.

To construct the probability distribution given the observed data set $\mathcal{D}_i^{\mathrm{obs}}$, we first consider the theoretical descriptions of the data at a point $\mathbf{x}_i$ and compute their ensemble average

$$\mathcal{D}(\mathbf{x}_i) = \bar{\mathcal{D}}(z_i)\left[1 + \delta\mathcal{D}(\mathbf{x}_i)\right] , \qquad \mu_i := \langle \mathcal{D}(\mathbf{x}_i)\rangle = \bar{\mathcal{D}}(z_i) , \tag{3.1}$$

where the ensemble average is performed over the hypersurface of fixed observed redshift (see [36] for the subtle gauge issues associated with the ensemble average in a usual coordinate). We now define the deviation $\Delta\mathcal{D}$ from the mean and its covariance $\mathbf{C}$:

$$\Delta\mathcal{D}(\mathbf{x}_i) := \mathcal{D}(\mathbf{x}_i) - \mu_i = \bar{\mathcal{D}}(z_i)\delta\mathcal{D}(\mathbf{x}_i) , \qquad \mathbf{C}_{ij} := \langle \Delta\mathcal{D}(\mathbf{x}_i)\Delta\mathcal{D}(\mathbf{x}_j)\rangle = \bar{\mathcal{D}}(z_i)\bar{\mathcal{D}}(z_j)\, \xi_{ij} , \tag{3.2}$$

where we defined the dimensionless two-point correlation function

$$\xi_{ij} := \langle \delta\mathcal{D}(\mathbf{x}_i)\delta\mathcal{D}(\mathbf{x}_j)\rangle . \tag{3.3}$$

Remember that the deviation $\Delta\mathcal{D}(\mathbf{x}_i)$ of the observed data from the mean includes the background quantity $\bar{\mathcal{D}}(z_i)$ and the perturbation $\delta\mathcal{D}(\mathbf{x}_i)$. For a Gaussian distribution, the two-point correlation function ξ_{ij} contains all the information. In real observations, the observed data set $\mathcal{D}_i^{\mathrm{obs}}$ is a set of numbers associated to the observed positions $\mathbf{x}_i$. The mean μ_i and the covariance $\mathbf{C}_{ij}$ are predictions of our chosen model at the observed positions $\mathbf{x}_i$.

Last, we need the inverse covariance $\mathbf{K}_{ij}$ to construct the probability distribution, defined by

$$\sum_j^N \mathbf{C}_{ij}\mathbf{K}_{jk} = \mathbf{I}_{ik} , \tag{3.4}$$

where $\mathbf{I}$ is an identity matrix. The Gaussian probability distribution of the observed data set $\mathcal{D}_i^{\rm obs}$ can now be written as

$$\mathcal{P} = \frac{1}{\left[(2\pi)^N \det \mathbf{C}\right]^{1/2}} \exp\left[-\frac{1}{2}\sum_{ij}^{N} \Delta\mathcal{D}_i \mathbf{K}_{ij} \Delta\mathcal{D}_j\right] . \tag{3.5}$$

It is the probability that we measure the observed data set given our theoretical predictions, and its logarithm is referred to as the likelihood $\mathcal{L} := -\ln\mathcal{P}$. With our cosmological model and its parameters, we predict the mean μ_i, the covariance matrix $\mathbf{C}_{ij}$, and its inverse $\mathbf{K}_{ij}$. As described, the likelihood analysis proceeds as follows: We first choose a set of cosmological parameters and predict μ_i, $\mathbf{C}_{ij}$ and $\mathbf{K}_{ij}$ to compute $\mathcal{P}$ or $\mathcal{L}$. We then explore the cosmological parameter space to maximize $\mathcal{P}$, given the observed data set.

3.2 Maximum cosmological information on the light cone

The probability distribution is maximized, only if we choose the best cosmological model and its parameters. Even after the maximum is found, however, the cosmological information of the observed data set $\mathcal{D}_i^{\rm obs}$ in a given survey is *not infinite*, but *finite*. Furthermore, given the survey volume, more information can be extracted, if we make more observations. Then the key question naturally arises: "*What is the maximum cosmological information content on the light cone (up to some redshift)?*" This question has not been properly addressed in literature.

To compute the maximum cosmological information contained in the light cone volume, we make a series of idealized assumptions: The observed data set is free of any systematic errors or measurement errors and is obtained from an *infinite* number ($N = \infty$) of galaxies within the survey boundary (no shot-noise contribution $n_g = \infty$). Of course, we need to know the correct cosmological model and parameters. Under these assumptions, the (discrete) observed data set $\mathcal{D}_i^{\rm obs}$ becomes a continuous field $\mathcal{D}(\mathbf{x})$ in Eq. (2.19), and the probability distribution in Eq. (3.5) can be trivially generalized to this idealized case by replacing the discrete sum with an integral. In this limit, the covariance $\mathbf{C}_{ij}$ and the correlation function ξ_{ij} are also naturally promoted to the continuous fields, and the inverse covariance is then subject to the continuous orthonormality condition, rather than the discrete one in Eq. (3.4):

$$(\mathbf{C}\,\mathbf{K})(\mathbf{x}_1,\mathbf{x}_2) = dz_1 d\Omega_1 \int dz' \int d\Omega'\, \xi(\mathbf{x}_1,\mathbf{x}')\zeta(\mathbf{x}',\mathbf{x}_2) = \delta^D(z_1-z_2)\delta^D(\Omega_1-\Omega_2)dz_1 d\Omega_1 , \tag{3.6}$$

where we replaced the Kronecker delta with the Dirac delta function, and the indices $i, j, \cdots$ are also replaced with continuous field variables. Note that both $\mathbf{C}$ and $\mathbf{K}$ are infinite-dimensional matrices. We also defined the "inverse" correlation function $\zeta(\mathbf{x}_1,\mathbf{x}_2)$ with $\hat{\bar{\mathcal{D}}}$ (not with $\bar{\mathcal{D}}$) as

$$\mathbf{K}_{12} := \mathbf{K}(\mathbf{x}_1,\mathbf{x}_2) := \left[\hat{\bar{\mathcal{D}}}(z_1)\hat{\bar{\mathcal{D}}}(z_2)\right]^{-1}\zeta(\mathbf{x}_1,\mathbf{x}_2) , \qquad\qquad \zeta_{12} := \zeta(\mathbf{x}_1,\mathbf{x}_2) , \tag{3.7}$$

where we defined a short-hand notation for the infinite dimensional matrices. From now on, we will work with the continuous fields to compute the maximum cosmological information contents on the light cone.

3.3 Decompositions and inverse covariance matrix

The observation on the light cone is made in terms of the observed redshift and angular position, and the radial information from the redshift represents not only the radial coordinate, but also the time along the past light cone. This rather trivial observation demands that we treat these two fundamental observables *differently*. While statistical isotropy implies that harmonic modes in angular direction are statistically independent, radial modes are not, even if we assume statistical homogeneity, since on the light cone they mix spatial and time directions.

First, we compute the inverse covariance in a single redshift bin in terms of spherical harmonics decomposition of the angular position, where the radial information is fixed or integrated out. Second, we compute the inverse covariance utilizing the full 3D information in terms of spherical Fourier decomposition.

3.3.1 Single redshift bin: Angular decomposition

As described in section 2.3, we may be interested in the observables in a single redshift bin such as the type-Ia supernovae at a given redshift and the CMB temperature anisotropies today. These observables are essentially a function of angular position alone, and the angular decomposition based on spherical harmonics provides a useful description.

In a single redshift bin, the observed data set is modeled in section 2.3 and summarized in Table 1 as

$$\mathcal{D}(\mathbf{x}) = \bar{D}(z)\frac{d\Omega}{4\pi}\left[1+\delta\mathcal{D}(\mathbf{x})\right] , \qquad \mu = \bar{D}(z)\frac{d\Omega}{4\pi} , \qquad \hat{\bar{\mathcal{D}}}(z) = \frac{\bar{D}(z)}{4\pi\Delta z} . \tag{3.8}$$

The deviation from the mean, the covariance matrix, and the inverse covariance matrix are then

$$\Delta\mathcal{D}(\mathbf{x}) = \bar{D}(z)\frac{d\Omega}{4\pi}\,\delta\mathcal{D}(\mathbf{x}) , \qquad \mathbf{C}_{12} = \bar{D}^2(z)\frac{d\Omega_1}{4\pi}\frac{d\Omega_2}{4\pi}\,\xi_{12} , \qquad \mathbf{K}_{12} = \left(\frac{4\pi\Delta z}{\bar{D}(z)}\right)^2\zeta_{12} , \tag{3.9}$$

where $\mathbf{x}_1 = (z, \hat{\mathbf{n}}_1)$ and $\mathbf{x}_2 = (z, \hat{\mathbf{n}}_2)$. The orthonormality relation in Eq. (3.6) can be integrated over the redshift width to derive the orthonormality relation in a single redshift bin

$$\int_z\int_{z'}(\mathbf{CK})_{12} = (\Delta z)^2 d\Omega_1\int d\Omega'\,\xi(\mathbf{x}_1,\mathbf{x}')\zeta(\mathbf{x}',\mathbf{x}_2) = \delta^D(\Omega_1-\Omega_2)d\Omega_1 . \tag{3.10}$$

To make further progress, we decompose the angular position $\hat{\mathbf{n}}$ of the observable fluctuation in terms of spherical harmonics Y_{lm} as

$$\delta\mathcal{D}(\mathbf{x}) := \sum_{lm} a_{lm}(z)Y_{lm}(\hat{\mathbf{n}}) , \qquad a_{lm}(z) \equiv \int d\Omega\, Y^*_{lm}(\hat{\mathbf{n}})\delta\mathcal{D}(\mathbf{x}) , \tag{3.11}$$

where a_{lm} is the angular coefficient that depends on the redshift. Using statistical isotropy, the angular power spectrum is

$$\langle a_{lm}a^*_{l'm'}\rangle = \delta_{ll'}\delta_{mm'}C_l , \tag{3.12}$$

and the two-point correlation function can be angular decomposed as

$$\xi_{12} = \langle\delta\mathcal{D}(\mathbf{x}_1)\delta\mathcal{D}(\mathbf{x}_2)\rangle = \sum_{lm} C_l Y_{lm}(\hat{\mathbf{n}}_1)Y^*_{lm}(\hat{\mathbf{n}}_2) = \sum_l \frac{2l+1}{4\pi}C_l P_l(\gamma_{12}) , \qquad \gamma_{12} := \hat{\mathbf{n}}_1\cdot\hat{\mathbf{n}}_2 , \tag{3.13}$$

where P_l is the Legendre polynomial of degree l. The inverse correlation function will be angular decomposed in the exactly same way, defining the "inverse" angular power spectrum $\tilde{C}_l$

$$\zeta_{12} := \sum_{lm}\tilde{C}_l Y_{lm}(\hat{\mathbf{n}}_1)Y^*_{lm}(\hat{\mathbf{n}}_2) = \sum_l \frac{2l+1}{4\pi}\tilde{C}_l P_l(\gamma_{12}) . \tag{3.14}$$

By using the orthonormality condition in Eq. (3.10), the inverse angular power spectrum can be readily obtained as

$$\tilde{C}_l = \frac{1}{C_l(\Delta z)^2} \,, \tag{3.15}$$

where we used

$$\sum_{lm} Y_{lm}(\hat{\mathbf{n}}_1) Y^*_{lm}(\hat{\mathbf{n}}_2) = \delta^D(\Omega_1 - \Omega_2) = \frac{1}{\sin\theta_1}\delta^D(\theta_1 - \theta_2)\delta^D(\phi_2 - \phi_2) \,. \tag{3.16}$$

Indeed, the inverse angular power spectrum $\tilde{C}_l$ is the inverse of the angular power spectrum C_l, and the inverse covariance is explicitly

$$\mathbf{K}_{12} = \frac{4\pi}{\bar{D}(z)^2} \sum_l \frac{2l+1}{C_l} P_l(\gamma_{12}) \,. \tag{3.17}$$

3.3.2 Light cone volume: Spherical Fourier decomposition

In addition to the angular position $\hat{\mathbf{n}}$, the cosmological observables are marked by their radial position in terms of the observed redshift z. With this additional radial dimension, we perform the spherical Fourier decomposition. As opposed to the standard Fourier analysis in a rectangular coordinate, the spherical Fourier analysis utilizes the eigenfunctions of the Helmholtz equation in a spherical coordinate, which naturally describes our fundamental observables $(z, \hat{\mathbf{n}})$. The spherical Fourier analysis was well developed [11–13] in galaxy clustering and applied to the baryonic acoustic oscillation [16], the three-dimensional weak lensing [55], and the relativistic effects [19]. Here we will adopt the formalism and notation convention in [19].

Given the three-dimensional positional information of the cosmological observables over the light cone volume, we apply the spherical Fourier analysis to compute the spherical power spectrum, and the fluctuations in the observable are now decomposed as

$$\delta\mathcal{D}(\mathbf{x}) := \sum_{lm} \int_0^\infty dk \sqrt{\frac{2}{\pi}}\, k j_l(k\bar{r}) Y_{lm}(\hat{\mathbf{n}}) s_{lm}(k) \,, \tag{3.18}$$

where $j_l(x)$ is the spherical Bessel function and the spherical Fourier coefficients are

$$s_{lm}(k) \equiv \int d\Omega \int d\bar{r}\, \bar{r}^2 \sqrt{\frac{2}{\pi}}\, k j_l(k\bar{r}) Y^*_{lm}(\hat{\mathbf{n}}) \delta\mathcal{D}(\mathbf{x}) \,, \tag{3.19}$$

where the radial integration is limited to the survey range. Here we have assumed a fiducial cosmological model to relate the measured redshift to a radial distance $\bar{r}$. Hence, both $\bar{r}$ and k are model dependent and not direct observables. The amplitude square of the spherical Fourier coefficients $s_{lm}(k)$ is the spherical power spectrum

$$\left\langle s_{lm}(k) s^*_{l'm'}(k') \right\rangle = \delta_{ll'}\delta_{mm'} S_l(k, k') \,, \tag{3.20}$$

where we assume statistical isotropy. The spherical power spectrum in the simplest case in Appendix B will be identical to the usual power spectrum $S_l(k, k') = \delta^D(k - k')P(k)$. Mind that the dimensions of the Fourier mode and its spherical power spectrum are somewhat different from the usual Fourier analysis:

$$[s_{lm}(k)] = L^2 \,, \qquad\qquad [S_l(k, k')] = L^4 \,. \tag{3.21}$$

In our spherical Fourier decomposition, we assumed that the sky coverage is the full sky. In real surveys, the sky coverage is never the full sky, but our assumption is fine, because while the low angular multipoles in real surveys are inevitably correlated, the high angular multipoles are independent even with an incomplete sky coverage. The radial Fourier modes would be independent if the survey volume is infinite and the fluctuations are time-translation invariant along the line-of-sight direction. However, none of these two are true, and we take into account that $S_l(k,k')$ is not diagonal since it mixes spatial and temporal information and our background universe is not time-translation invariant.

Given the spherical Fourier decomposition, we use the orthonormality condition to derive the inverse covariance $\mathbf{K}_{ij}$ or the inverse correlation function ζ_{ij}. We first compute the two-point correlation function in terms of the spherical power spectrum $S_l(k,k')$:

$$\xi_{12} = \langle \delta\mathcal{D}(\mathbf{x}_1)\delta\mathcal{D}(\mathbf{x}_2)\rangle = 4\pi\sum_{lm}\int dk\int dk'\,\frac{kk'}{2\pi^2}S_l(k,k')j_l(k\bar{r}_1)j_l(k'\bar{r}_2)Y_{lm}(\hat{\mathbf{n}}_1)Y^*_{lm}(\hat{\mathbf{n}}_2)\;, \tag{3.22}$$

and write the inverse correlation function in the similar way

$$\zeta_{12} := 4\pi\sum_{lm}\int dk\int dk'\,\frac{kk'}{2\pi^2}\tilde{S}_l(k,k')j_l(k\bar{r}_1)j_l(k'\bar{r}_2)Y_{lm}(\hat{\mathbf{n}}_1)Y^*_{lm}(\hat{\mathbf{n}}_2)\;, \tag{3.23}$$

defining the "inverse" spherical power spectrum $\tilde{S}_l(k,k')$. Due to the rotational symmetry, both the correlation functions are only a function of angular multipole l, but we keep the explicit dependence on $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ for later convenience. Using the spherical Fourier decomposition of ξ_{12} and ζ_{12} and integrating over $d\Omega'$, the orthonormality condition in Eq. (3.6) can be written as

$$\begin{aligned}\delta^D(z_1-z_2)\delta^D(\Omega_1-\Omega_2) \;=\; &(4\pi)^2\sum_{lm}Y_{lm}(\hat{\mathbf{n}}_1)Y^*_{lm}(\hat{\mathbf{n}}_2)\int dk_1\int dk_2\int dk_1'\int dk_2'\,\frac{k_1k_2}{2\pi^2}\,\frac{k_1'k_2'}{2\pi^2}\\ &\times S_l(k_1,k_2)\tilde{S}_l(k_1',k_2')j_l(k_1\bar{r}_1)j_l(k_2'\bar{r}_2)\mathcal{F}_l(k_2,k_1')\;,\end{aligned} \tag{3.24}$$

where we defined the Fourier angular kernel

$$\mathcal{F}_l(k_1,k_2) := \int dz\; j_l(k_1\bar{r})j_l(k_2\bar{r})\;, \tag{3.25}$$

and the integration range is again limited to the survey range. Since the inverse spherical power spectrum is part of the integral equation, we make a series of manipulations to pull out the inverse spherical power spectrum from the integral by using the identities associated with the Dirac delta function. First, we multiply the orthonormality equation by $Y_{LM}(\hat{\mathbf{n}}_2)$ and integrate over $d\Omega_2$ to obtain

$$\begin{aligned}&(4\pi)^2\int dk_1\int dk_2\int dk_1'\int dk_2'\,\frac{k_1k_2}{2\pi^2}\,\frac{k_1'k_2'}{2\pi^2}\\ &\qquad\times S_l(k_1,k_2)\tilde{S}_l(k_1',k_2')j_l(k_1\bar{r}_1)j_l(k_2'\bar{r}_2)\mathcal{F}_l(k_2,k_1') = \delta^D(z_1-z_2)\;,\end{aligned} \tag{3.26}$$

where we re-labeled the angular index L by l. It is noted that the orthonormality condition after the angular integration becomes independent of angular multipoles l, or identical for all angular multipoles. To further simplify the condition, we use the mathematical identity of spherical Bessel functions (independent of the survey depth)

$$\int_0^\infty d\bar{r}\;\bar{r}^2 j_l(k_1\bar{r})j_l(k_2\bar{r}) = \frac{\pi}{2k_1k_2}\delta^D(k_1-k_2)\;, \tag{3.27}$$

and integrate over $d\bar{r}_2$ after multiplying by $\bar{r}_2^2 j_l(k_A\bar{r}_2)$ to derive

$$4\pi\int dk_1\int dk_2\int dk_1'\,\frac{k_1k_2}{2\pi^2}\,\frac{k_1'}{k_A}S_l(k_1,k_2)\tilde{S}_l(k_1',k_A)j_l(k_1\bar{r}_i)\mathcal{F}_l(k_2,k_1')=\left(\frac{\bar{r}^2}{H}\right)_{z_1}j_l(k_A\bar{r}_1)\,. \tag{3.28}$$

Finally, we use the identity of the spherical Bessel function one more time by multiplying by $j_l(k_B\bar{r}_1)$ and integrating over dz_1, and by re-arranging the equation, we arrive at the closed equation for the inverse spherical power spectrum

$$\left(\frac{2}{\pi}\right)^2\int dk_1'\,k_1'k_B\tilde{S}_l(k_1',k_A)\times\int dk_1\int dk_2\,k_1k_2S_l(k_1,k_2)\mathcal{F}_l(k_2,k_1')\mathcal{F}_l(k_B,k_1)=\delta^D(k_A-k_B)\,. \tag{3.29}$$

This integral equation is multiplications of infinite-dimensional matrices, resulting in the identity matrix:

$$\tilde{\mathbf{S}}_l\left(\mathbf{F}_l\mathbf{S}_l\mathbf{F}_l\right)=\mathbf{I}\,,\qquad\qquad \tilde{\mathbf{S}}_l=\left(\mathbf{F}_l\mathbf{S}_l\mathbf{F}_l\right)^{-1}\,, \tag{3.30}$$

where we used the boldface letters for the matrices

$$(\tilde{\mathbf{S}}_l)_{12}=\tilde{S}_l(k_1,k_2)k_2\,,\qquad (\mathbf{S}_l)_{12}=S_l(k_1,k_2)k_2\,,\qquad (\mathbf{F}_l)_{12}=\frac{2}{\pi}\,\mathcal{F}_l(k_1,k_2)k_2\,. \tag{3.31}$$

The inverse spherical power spectrum should be obtained by inverting this matrix equation. In section 5.5, we derive the inverse spherical power spectrum $\tilde{S}_l(k,k')\propto P^{-1}(k)\delta(k-k')$ in Eq. (5.27) for a small survey volume, where the flat-sky approximation and no time evolution can be adopted. Therefore, the inverse covariance in the light cone volume is

$$\mathbf{K}_{12}=\frac{4\pi}{\hat{\bar{\mathcal{D}}}(z_1)\hat{\bar{\mathcal{D}}}(z_2)}\sum_l\frac{2l+1}{4\pi}P_l(\gamma_{12})\int dk\int dk'\,\frac{kk'}{2\pi^2}\tilde{S}_l(k,k')j_l(k\bar{r}_1)j_l(k'\bar{r}_2)\,. \tag{3.32}$$

3.3.3 Projected observables: Angular decomposition

The observables in a single redshift bin in section 2.4 were angular decomposed in section 3.3.1, but they can also be described in terms of spherical Fourier decomposition. The relation of the spherical power spectrum to the angular power spectrum can be read off as

$$a_{lm}(z)=\int_0^\infty dk\sqrt{\frac{2}{\pi}}kj_l(k\bar{r}_z)s_{lm}(k)\,,\qquad\qquad C_l(z)=4\pi\int dk\int dk'\,\frac{kk'}{2\pi^2}S_l(k,k')j_l(k\bar{r}_z)j_l(k'\bar{r}_z)\,. \tag{3.33}$$

It is evident that the inverse covariance in Eq. (3.32) over the light cone volume is not the inverse covariance in Eq. (3.17), even when the survey volume is restricted to the sub-volume of a single redshift bin, i.e., if more data beyond the single redshift bin are available, we have to account for the radial correlation and more information is available, even when we restrict our analysis to a single redshift bin.

The projected observables such as the weak lensing observables and the projected galaxy number density are somewhat different from what we derived for a single redshift bin in section 3.3.1, as the three-dimensional spherical power spectrum is integrated along the line-of-sight direction. Given the projected observable quantity $\mathcal{D}^{\rm pro}(\hat{\mathbf{n}})$ in Eq. (2.33), we construct the covariance

$$\mathbf{C}_{12}=d\Omega_1d\Omega_2\int dz_1\left(\frac{1}{\bar{N}}\frac{d\bar{N}}{dzd\Omega}\right)_{z_1}\int dz_2\left(\frac{1}{\bar{N}}\frac{d\bar{N}}{dzd\Omega}\right)_{z_2}\langle\delta\mathcal{D}(\mathbf{x}_1)\delta\mathcal{D}(\mathbf{x}_2)\rangle=:\xi_{12}d\Omega_1d\Omega_2\,, \tag{3.34}$$

where the ensemble average in the integrand is the (three-dimensional) two-point correlation function in Eq. (3.22) and we define the (projected) angular two-point correlation ξ_{12}, in comparison to Eq. (3.9). The angular two-point correlation function can be further decomposed as in Eq. (3.13), and the angular power spectrum of the projected observable $\mathcal{D}^{\rm pro}$ is then

$$C_l = 4\pi \int dk \int dk' \, \frac{kk'}{2\pi^2} S_l(k,k') \left[\int dz_1 \left(\frac{1}{\bar{N}} \frac{d\bar{N}}{dzd\Omega}\right) j_l(k\bar{r}_1)\right] \times \left[\int dz_2 \left(\frac{1}{\bar{N}} \frac{d\bar{N}}{dzd\Omega}\right) j_l(k'\bar{r}_2)\right] . \tag{3.35}$$

The angular power spectrum in Eq. (3.33) is recovered, if the redshift distribution is confined to a single redshift bin. Similarly, the inverse covariance is defined in terms of the inverse correlation function ζ_{12}

$$\mathbf{K}_{12} := \zeta_{12} = \sum_l \frac{2l+1}{4\pi} \tilde{C}_l P_l(\gamma_{12}) \,, \tag{3.36}$$

and the inverse angular power spectrum is

$$\tilde{C}_l = \frac{1}{C_l} \,, \tag{3.37}$$

in similarity to Eq. (3.14).

4 Cosmological Information Contents on the Light Cone

As discussed in section 3.2, we are interested in the maximum possible cosmological information derivable from a given observable measured in the light cone volume. To compute these cosmological information contents, we employ the Fisher information formalism. The Fisher information technique has been well developed for galaxy clustering and CMB analysis (see, e.g., [53, 54]). However, since cosmological observables are mapped in terms of the redshift and the angular position, the standard flat-sky or Euclidean descriptions in the Fisher analysis represent only the approximation to the real observations. Our Fisher matrix analysis based on the redshift and the angular position provides the most accurate description of the cosmological information contents in the light cone volume, and further insights can be gained in connection to the standard analysis in the limiting cases, where the survey volume is small and narrow.

Moreover, our formalism in section 2 reveals that cosmological observables come with additional fluctuations associated with the light propagation and the galaxy number counts. For instance, observations of supernova type Ia are used to measure the background luminosity distance $\bar{\mathcal{D}}_L(z)$. However, what we measure is indeed the full luminosity distance $\mathcal{D}_L(z) = \bar{\mathcal{D}}_L(z)(1+\delta\mathcal{D}_L)$, including both the background and the fluctuations, as supernovae can go off only in some host galaxies (not in random places) and the light propagation is thereafter affected by the fluctuations between the source and the observer. Consequently, even if we could measure an infinite number of supernovae in a single redshift bin, the uncertainty in our luminosity distance measurements is limited by the cosmic variance, and the maximum cosmological information is certainly *not* infinite, even in this idealized situation. We will compute this maximum cosmological information contents on the light cone.

First, we briefly review the Fisher information technique, and we then proceed to compute the cosmological information contents in a single redshift bin and on the light cone.

4.1 Fisher information matrix

The Fisher information formalism has been well developed and applied in cosmology (see, e.g., [53, 54]), and we provide a short description to fix our notational convention.

To compute the derivatives of the log likelihood $\mathcal{L} = -\ln \mathcal{P}$ around the maximum with respect to the model parameters p_μ, we first construct the data matrix out of the individual data $\mathcal{D}_i^{\rm obs}$ and the model prediction μ_i

$$\mathbf{D} := (\boldsymbol{\mathcal{D}} - \boldsymbol{\mu})(\boldsymbol{\mathcal{D}} - \boldsymbol{\mu})^{\rm t} = \Delta\mathcal{D}_i \Delta\mathcal{D}_j \,, \tag{4.1}$$

where we used the boldface letter to indicate that the quantities are vectors and matrices. Choosing cosmological parameters such that $\boldsymbol{\mu}$ is the best fit to the data $\boldsymbol{\mathcal{D}}$, the ensemble average of first derivative of the data matrix vanishes,

$$\left\langle \frac{\partial}{\partial p_\mu} \mathbf{D} \right\rangle = 0 \,, \tag{4.2}$$

but the second derivative of the data matrix carries important information about the variation of the mean with respect to the model parameters

$$\mathbf{M}_{\mu\nu} := \left\langle \frac{\partial^2 \mathbf{D}}{\partial p_\mu \partial p_\nu} \right\rangle = \frac{\partial \boldsymbol{\mu}}{\partial p_\mu} \frac{\partial \boldsymbol{\mu}^{\rm t}}{\partial p_\nu} + \frac{\partial \boldsymbol{\mu}}{\partial p_\nu} \frac{\partial \boldsymbol{\mu}^{\rm t}}{\partial p_\mu} \,. \tag{4.3}$$

Given the Gaussian probability distribution in Eq. (3.5), we compute the log-likelihood

$$-2 \ln \mathcal{P} = \ln \det \mathbf{C} + \sum_{ij} \Delta\mathcal{D}_i \mathbf{K}_{ij} \Delta\mathcal{D}_j + N \ln 2\pi \,, \tag{4.4}$$

and the Fisher information matrix is then

$$F_{\mu\nu} := \left\langle -\frac{\partial^2 \ln \mathcal{P}}{\partial p_\mu \partial p_\nu} \right\rangle = \frac{1}{2} \mathrm{Tr} \left[\left(\mathbf{K} \frac{\partial}{\partial p_\mu} \mathbf{C} \right) \left(\mathbf{K} \frac{\partial}{\partial p_\nu} \mathbf{C} \right) + \mathbf{K}\mathbf{M}_{\mu\nu} \right] \,. \tag{4.5}$$

To leading order, the constant likelihood surface in parameter space is determined by two distinct contributions: the variations of the mean and the two-point correlation, given the inverse covariance. The inverse of the Fisher matrix is the optimistic forecast for the parameter estimation, representing the cosmological information in the observed data set. We again generalize this Fisher matrix formalism for the (discrete) observed data set $\mathcal{D}_i^{\rm obs}$ to a continuous field $\mathcal{D}(\mathbf{x})$.

4.2 Single redshift bin and projected observables

Cosmological observables in a single redshift bin or the projected observables are decomposed in terms of spherical harmonics, and we have derived the inverse covariance matrix in section 3.3. We compute the Fisher information matrix in Eq. (4.5), using the angular decomposition.

Given the cosmological observable $\mathcal{D}(\mathbf{x})$ and its mean μ in Eq. (3.8) in a single redshift bin, we first compute the variation of the mean with respect to the model parameters

$$(\mathbf{M}_{\mu\nu})_{12} = 2 \, \frac{\partial \bar{D}(z)}{\partial p_\mu} \frac{\partial \bar{D}(z)}{\partial p_\nu} \frac{d\Omega_1}{4\pi} \frac{d\Omega_2}{4\pi} \,, \tag{4.6}$$

and its product with the inverse covariance is

$$\frac{1}{2} (\mathbf{K}\mathbf{M}_{\mu\nu})_{13} = (\Delta z)^2 \, \frac{\partial \ln \bar{D}(z)}{\partial p_\mu} \frac{\partial \ln \bar{D}(z)}{\partial p_\nu} \, d\Omega_3 \int d\Omega_2 \, \zeta_{12} \,. \tag{4.7}$$

Therefore, the first contribution to the Fisher information matrix can be obtained by taking the trace and using the angular power spectrum decomposition in Eq. (3.14) as

$$\frac{1}{2}\mathrm{Tr}\Big[\mathbf{K}\mathbf{M}_{\mu\nu}\Big] \equiv (\Delta z)^2 \frac{\partial \ln \bar{D}(z)}{\partial p_\mu}\frac{\partial \ln \bar{D}(z)}{\partial p_\nu}\int d\Omega_1 \int d\Omega_2 \, \zeta_{12} = \frac{4\pi}{C_0}\left(\frac{\partial \ln \bar{D}(z)}{\partial p_\mu}\right)\left(\frac{\partial \ln \bar{D}(z)}{\partial p_\nu}\right) , \tag{4.8}$$

where C_0 is the monopole of the angular power spectrum C_l and happens to be the cosmic variance of $\bar{D}(z)$ [43]. Here we see that C_0 limits indeed the measurements of $\bar{D}(z)$ from the Fisher matrix. We emphasize that the cosmic variance of the background quantity in the measurements is *not* appreciated in literature. The reason is that one often equates the angular average with the ensemble average, such that the monopole contribution $\delta\mathcal{D}_0$ is set zero, but as shown in Eq. (2.24) this is *not* the case. The equality holds, only when we perform the angular average over many different observer positions, assuming the Ergodic hypothesis [43].

The other contribution to the Fisher information matrix is the variation of the covariance matrix with respect to the model parameters. The covariance matrix in a single redshift bin is given in Eq. (3.9), and its product with the inverse covariance can be obtained as

$$\begin{aligned}\left(\mathbf{K}\frac{\partial}{\partial p_\mu}\mathbf{C}\right)_{13} &\equiv \frac{(\Delta z)^2}{\bar{D}^2(z)}\int d\Omega_2\, \zeta_{12}\frac{\partial}{\partial p_\mu}\left[\bar{D}^2(z)\xi_{23}\right] d\Omega_3 \\ &= \sum_{lm} Y_{lm}(\hat{\mathbf{n}}_1)\frac{\partial}{\partial p_\mu}\Big(\ln\left[\bar{D}^2(z)C_l\right]\Big)Y^*_{lm}(\hat{\mathbf{n}}_3)d\Omega_3 \, ,\end{aligned} \tag{4.9}$$

where we used the angular power spectrum decomposition for ξ_{12} and ζ_{12}. By multiplying the same product with parameter p_ν and taking the trace of the product, we derive the covariance contribution to the Fisher information matrix

$$\frac{1}{2}\mathrm{Tr}\left[\left(\mathbf{K}\frac{\partial}{\partial p_\mu}\mathbf{C}\right)\left(\mathbf{K}\frac{\partial}{\partial p_\nu}\mathbf{C}\right)\right] = \sum_l \frac{2l+1}{2}\frac{\partial}{\partial p_\mu}\Big(\ln \bar{D}^2(z)C_l\Big)\frac{\partial}{\partial p_\nu}\Big(\ln \bar{D}^2(z)C_l\Big) \, . \tag{4.10}$$

The covariance $\mathbf{C}$ of the cosmological observable $\mathcal{D}(\mathbf{x})$ contains the extra information about the observable, fully characterized by the two-point correlation function $\xi(\mathbf{x}_1, \mathbf{x}_2)$ under the assumption of Gaussianity. In a single redshift bin, this information is better represented by the angular power spectrum C_l, and we recover the standard expression, often phrased as follows: The angular power spectrum measurements are limited as we can only sample $2l+1$ independent components of a_{lm}. This statement is true, but with one subtle caveat that we *cannot* directly measure $\delta\mathcal{D}(\mathbf{x})$ or its angular component a_{lm} in Eq. (2.25). What we can measure is the full cosmological observable $\mathcal{D}(\mathbf{x})$ that includes both the background $\bar{\mathcal{D}}(z)$ and the perturbation $\delta\mathcal{D}(\mathbf{x})$, such that our observed mean is limited by C_0 and our observed power spectrum is limited by C_l but only through the combination $\bar{D}^2(z)C_l$.

The full Fisher information matrix in a single redshift bin is then

$$F_{\mu\nu} = \frac{4\pi}{C_0}\left(\frac{\partial \ln \bar{D}(z)}{\partial p_\mu}\right)\left(\frac{\partial \ln \bar{D}(z)}{\partial p_\nu}\right) + \sum_l \frac{2l+1}{2}\frac{\partial}{\partial p_\mu}\Big(\ln \bar{D}^2(z)C_l\Big)\frac{\partial}{\partial p_\nu}\Big(\ln \bar{D}^2(z)C_l\Big) \, . \tag{4.11}$$

It represents the maximum information contained in the observed data set $\mathcal{D}_i^{\mathrm{obs}}$ in a single redshift bin under the assumption that the underlying fluctuations are Gaussian distributed. This result has been well-known in literature without the background part $\bar{D}(z)$ (see, however, [53]). For the projected observables in section 2.4, where the observables vanish in the background, the Fisher information matrix contains only the covariance part

$$F_{\mu\nu} = \sum_l \frac{2l+1}{2} \left(\frac{\partial}{\partial p_\mu} \ln C_l \right) \left(\frac{\partial}{\partial p_\nu} \ln C_l \right) , \tag{4.12}$$

where the angular power spectrum C_l of the projected observables is given in Eq. (3.35).

4.3 Observations of the three-dimensional light cone volume

With the radial information given by the observed redshift, cosmological observables in a light cone volume are decomposed in terms of spherical harmonics and spherical Bessel functions, and we have derived the inverse covariance matrix in section 3.3.2. Here we compute the Fisher information matrix in Eq. (4.5), using the spherical Fourier decomposition, and this provides the most accurate description of the cosmological information contents on a light cone.

The calculations on the light cone proceed in a similar way to those in a single redshift bin in section 4.2. We compute the matrix multiplications for the Fisher information matrix and perform the spherical Fourier decomposition, instead of the angular decomposition used in section 4.2. The variation of the mean with respect to the model parameters is

$$(\mathbf{M}_{\mu\nu})_{12} = \left[\frac{\partial}{\partial p_\mu} \hat{\bar{\mathcal{D}}}(z_1) \frac{\partial}{\partial p_\nu} \hat{\bar{\mathcal{D}}}(z_2) + \frac{\partial}{\partial p_\nu} \hat{\bar{\mathcal{D}}}(z_1) \frac{\partial}{\partial p_\mu} \hat{\bar{\mathcal{D}}}(z_2) \right] dz_1 d\Omega_1 dz_2 d\Omega_2 , \tag{4.13}$$

and its product with the inverse covariance is

$$\frac{1}{2} (\mathbf{K}\mathbf{M}_{\mu\nu})_{13} = \frac{dz_3 d\Omega_3}{2\hat{\bar{\mathcal{D}}}(z_1)} \int dz_2 \int d\Omega_2 \, \zeta_{12} \left[\frac{\partial}{\partial p_\mu} \ln \hat{\bar{\mathcal{D}}}(z_2) \frac{\partial}{\partial p_\nu} \hat{\bar{\mathcal{D}}}(z_3) + \frac{\partial}{\partial p_\nu} \ln \hat{\bar{\mathcal{D}}}(z_2) \frac{\partial}{\partial p_\mu} \hat{\bar{\mathcal{D}}}(z_3) \right] . \tag{4.14}$$

Taking the trace of the product and using the spherical Fourier decomposition in Eq. (3.23), we derive the first contribution to the Fisher information matrix in the three-dimensional light cone volume

$$\begin{aligned} \frac{1}{2} \mathrm{Tr} \Big[\mathbf{K}\mathbf{M}_{\mu\nu} \Big] &\equiv \int dz_1 \int d\Omega_1 \int dz_2 \int d\Omega_2 \, \frac{1}{2} \zeta_{12} \left[\left(\frac{\partial}{\partial p_\mu} \ln \hat{\bar{\mathcal{D}}}(z_1) \right) \left(\frac{\partial}{\partial p_\nu} \ln \hat{\bar{\mathcal{D}}}(z_2) \right) + (1 \leftrightarrow 2) \right] \\ &= (4\pi)^2 \int dk \int dk' \, \frac{kk'}{2\pi^2} \tilde{S}_0(k,k') \mathcal{G}_\mu(k) \mathcal{G}_\nu(k') , \end{aligned} \tag{4.15}$$

where we used the spherical Fourier decomposition and defined the Fourier kernel

$$\mathcal{G}_\mu(k) := \int dz \, j_0(k\bar{r}) \frac{\partial}{\partial p_\mu} \ln \hat{\bar{\mathcal{D}}}(z) . \tag{4.16}$$

Note that $\bar{r}$ depends on the cosmological model.

Compared to Eq. (4.8), the contribution of the variation in the mean to the Fisher matrix information takes the similar structure: the variation of the mean value with respect to the model parameters is limited by the inverse monopole power spectrum $\tilde{S}_0(k,k')$. The Fisher information matrix shows that there exists the cosmic variance in the background quantity $\hat{\bar{\mathcal{D}}}(z)$, again set by the monopole ($l=0$), but integrated over different Fourier modes in the light cone volume.

We then move to compute the variation of the covariance matrix with respect to the model parameters. First, we compute the product of the inverse covariance and the derivative of the covariance matrix

$$\left(\mathbf{K}\frac{\partial}{\partial p_\mu}\mathbf{C}\right)_{13} = \frac{dz_3 d\Omega_3}{\hat{\bar{\mathcal{D}}}(z_1)}\int dz_2 \int d\Omega_2 \left[\frac{\zeta_{12}}{\hat{\bar{\mathcal{D}}}(z_2)}\right]\frac{\partial}{\partial p_\mu}\left[\hat{\bar{\mathcal{D}}}(z_2)\hat{\bar{\mathcal{D}}}(z_3)\xi_{23}\right] \tag{4.17}$$
$$= (4\pi)^2 \sum_{lm} \frac{Y_{lm}(\hat{\mathbf{n}}_1)Y^*_{lm}(\hat{\mathbf{n}}_3)}{\hat{\bar{\mathcal{D}}}(z_1)} dz_3 d\Omega_3 \int dk_1 \int dk_2 \frac{k_1k_2}{2\pi^2}\tilde{S}_l(k_1,k_2) j_l(k_1\bar{r}_1)$$
$$\times \int dz_2 \frac{j_l(k_2\bar{r}_2)}{\hat{\bar{\mathcal{D}}}(z_2)} \int dk_1' \int dk_2' \frac{k_1'k_2'}{2\pi^2}\frac{\partial}{\partial p_\mu}\left[\hat{\bar{\mathcal{D}}}(z_2)\hat{\bar{\mathcal{D}}}(z_3) j_l(k_1'\bar{r}_2) j_l(k_2'\bar{r}_3) S_l(k_1',k_2')\right] ,$$

where we integrated over $d\Omega_2$. By repeating the calculation with the parameter p_ν and taking the trace of the product, we obtain the other contribution to the Fisher information matrix

$$\frac{1}{2}\mathrm{Tr}\left[\left(\mathbf{K}\frac{\partial}{\partial p_\mu}\mathbf{C}\right)\left(\mathbf{K}\frac{\partial}{\partial p_\nu}\mathbf{C}\right)\right] = (4\pi)^4 \sum_l \frac{2l+1}{2}\int dz_1 \int dz_2 \int dz_3 \int dz_4 \tag{4.18}$$
$$\times \int dk_1 \int dk_2 \int dk_3 \int dk_4 \int dk_1' \int dk_2' \int dk_3' \int dk_4' \frac{k_1k_2}{2\pi^2}\frac{k_3k_4}{2\pi^2}\frac{k_1'k_2'}{2\pi^2}\frac{k_3'k_4'}{2\pi^2}$$
$$\times \frac{\tilde{S}_l(k_1,k_2)}{\hat{\bar{\mathcal{D}}}(z_1)\hat{\bar{\mathcal{D}}}(z_2)} j_l(k_1\bar{r}_1) j_l(k_2\bar{r}_2)\frac{\partial}{\partial p_\mu}\left[\hat{\bar{\mathcal{D}}}(z_2)\hat{\bar{\mathcal{D}}}(z_3) j_l(k_1'\bar{r}_2) j_l(k_2'\bar{r}_3) S_l(k_1',k_2')\right]$$
$$\times \frac{\tilde{S}_l(k_3,k_4)}{\hat{\bar{\mathcal{D}}}(z_3)\hat{\bar{\mathcal{D}}}(z_4)} j_l(k_3\bar{r}_3) j_l(k_4\bar{r}_4)\frac{\partial}{\partial p_\nu}\left[\hat{\bar{\mathcal{D}}}(z_4)\hat{\bar{\mathcal{D}}}(z_1) j_l(k_3'\bar{r}_4) j_l(k_4'\bar{r}_1) S_l(k_3',k_4')\right] .$$

To simplify the expression, we define a series of Fourier kernels, in addition to the Fourier angular kernel $\mathcal{F}_l(k_1,k_2)$ in Eq. (3.25):

$$\mathcal{F}_l(k_1,k_2) := \int dz\; j_l(k_1\bar{r}) j_l(k_2\bar{r}) =: \mathcal{F}_l^{12} , \tag{4.19}$$
$$\mathcal{H}_{l,\mu}(k_1,k_2) := \int dz\; j_l(k_1\bar{r}) j_l(k_2\bar{r}) \frac{\partial}{\partial p_\mu}\ln \hat{\bar{\mathcal{D}}}(z) =: \mathcal{H}_{l,\mu}^{12} , \tag{4.20}$$
$$\mathcal{N}_{l,\mu}(k_1,k_2) := \int dz\; j_l(k_1\bar{r}) \frac{\partial}{\partial p_\mu} j_l(k_2\bar{r}) =: \mathcal{N}_{l,\mu}^{12} , \tag{4.21}$$

where we used the super-scripts to simplify the arguments and the kernels $\mathcal{F}_l(k_1,k_2)$ and $\mathcal{H}_{l,i}(k_1,k_2)$ are symmetric in their arguments, but $\mathcal{N}_{l,\mu}(k_1,k_2)$ is not. Expanding the derivatives in Eq. (4.18) and integrating over the redshift, we derive

$$\frac{1}{2}\mathrm{Tr}\left[\left(\mathbf{K}\frac{\partial}{\partial p_\mu}\mathbf{C}\right)\left(\mathbf{K}\frac{\partial}{\partial p_\nu}\mathbf{C}\right)\right] = \left(\frac{2}{\pi}\right)^4 \sum_l \frac{2l+1}{2}\left(\prod_{i=1}^4 \int dk_i\; k_i\right)\left(\prod_{j=1}^4 \int dk_j'\; k_j'\right)\tilde{S}_l(k_1,k_2)\tilde{S}_l(k_3,k_4)$$
$$\times\left[\mathcal{F}_l^{21'}\left(\mathcal{H}_{l,\mu}^{32'}+\mathcal{N}_{l,\mu}^{32'}\right)S_l(k_1',k_2') + \mathcal{F}_l^{32'}\left(\mathcal{H}_{l,\mu}^{21'}+\mathcal{N}_{l,\mu}^{21'}\right)S_l(k_1',k_2') + \mathcal{F}_l^{21'}\mathcal{F}_l^{32'}\frac{\partial}{\partial p_\mu}S_l(k_1',k_2')\right]$$
$$\times\left[\mathcal{F}_l^{43'}\left(\mathcal{H}_{l,\nu}^{14'}+\mathcal{N}_{l,\nu}^{14'}\right)S_l(k_3',k_4') + \mathcal{F}_l^{14'}\left(\mathcal{H}_{l,\nu}^{43'}+\mathcal{N}_{l,\nu}^{43'}\right)S_l(k_3',k_4') + \mathcal{F}_l^{43'}\mathcal{F}_l^{14'}\frac{\partial}{\partial p_\nu}S_l(k_3',k_4')\right] . \tag{4.22}$$

Compared to Eq. (4.10), the contribution of the covariance to the Fisher information matrix is substantially more complicated, as it involves the three-dimensional spherical Fourier decomposition. However, the structure is similar in a sense that the three-dimensional fluctuations are correlated not only in angular directions, but also in radial direction, so that the measurements are limited by the cosmic variance given by the inverse spherical power spectrum $\tilde{S}_l(k,k')$. Furthermore, it clearly shows that the cosmological information is contained not only in the spherical power spectrum $S_l(k,k')$, but also in the angular diameter distance $\bar{r}$ and the background mean $\hat{\bar{\mathcal{D}}}(z)$ through the Fourier kernels $\mathcal{H}_{l,\mu}$ and $\mathcal{N}_{l,\mu}$.

Adding the two contributions, the full Fisher information matrix on the light cone can be written as

$$
\begin{aligned}
F_{\mu\nu} &= (4\pi)^2 \int dk \int dk' \, \frac{kk'}{2\pi^2} \tilde{S}_0(k,k') \mathcal{G}_\mu(k) \mathcal{G}_\nu(k') \\
&+ \left(\frac{2}{\pi}\right)^4 \sum_l \frac{2l+1}{2} \left(\prod_{i=1}^{4} \int dk_i \; k_i\right) \left(\prod_{j=1}^{4} \int dk'_j \; k'_j\right) \tilde{S}_l(k_1,k_2) \tilde{S}_l(k_3,k_4) \\
&\times \left[\mathcal{F}_l^{21'} \left(\mathcal{H}_{l,\mu}^{32'} + \mathcal{N}_{l,\mu}^{32'}\right) S_l(k'_1,k'_2) + \mathcal{F}_l^{32'} \left(\mathcal{H}_{l,\mu}^{21'} + \mathcal{N}_{l,\mu}^{21'}\right) S_l(k'_1,k'_2) + \mathcal{F}_l^{21'} \mathcal{F}_l^{32'} \frac{\partial}{\partial p_\mu} S_l(k'_1,k'_2)\right] \\
&\times \left[\mathcal{F}_l^{43'} \left(\mathcal{H}_{l,\nu}^{14'} + \mathcal{N}_{l,\nu}^{14'}\right) S_l(k'_3,k'_4) + \mathcal{F}_l^{14'} \left(\mathcal{H}_{l,\nu}^{43'} + \mathcal{N}_{l,\nu}^{43'}\right) S_l(k'_3,k'_4) + \mathcal{F}_l^{43'} \mathcal{F}_l^{14'} \frac{\partial}{\partial p_\nu} S_l(k'_3,k'_4)\right] .
\end{aligned}
\tag{4.23}
$$

This Fisher matrix represents the maximum cosmological information derivable from the observable in a light cone volume, again under the assumption that the underlying fluctuations are Gaussian distributed (see section 3.2). While the formal equation (4.5) for the Fisher information matrix was well known, this equation (4.23) in a three-dimensional light-cone volume is derived for the first time in this work.

Neglecting the terms $\mathcal{H}_{l,\nu}$ and $\mathcal{N}_{l,\nu}$ and using Limber approximation for the integrals over redshift (i.e. neglecting correlations at different redshifts) we obtain the well known results in 3D Fourier space. We shall show this in more details when treating examples in Section 5.

Before we terminate this section, we comment on the model-dependence of the power spectrum analysis (both the traditional and the spherical). The analysis involves a conversion of the observed galaxy position $\mathbf{x} = (z, \hat{\mathbf{n}})$ into the comoving distance $\bar{r}$, in which we need a prior cosmological model. In principle, this poses *no* problem, as the data processing in this case is part of the likelihood analysis, in which the raw observed data is re-processed for each set of cosmological parameters and compared to the theoretical predictions (see, e.g., [53, 56]). In practice, however, it is computationally more expensive to re-process the raw data for each likelihood ladder and hence a simple approximation is typically adopted (see, e.g., [57, 58]). In Appendix A, we present an alternative to the power spectrum analysis based on $\bar{r}$ by using the observed redshift itself as a dimensionless radial coordinate, and in this way the observed raw data can be processed only once in a model independent way.

5 Cosmological Applications

We apply our formalism to five different cosmological observables to compute the maximum cosmological information contents derivable from such observations in the idealized case described in section 3.2. Here we make a series of approximations and derive rough analytical estimates of the maximum cosmological information contents. A detailed analysis will require more extensive numerical investigations beyond our current scope.

5.1 Luminosity distance measurements in a single redshift bin with infinite number of supernovae

Consider observations of luminous type-Ia supernovae in a single redshift bin in a future survey, where a large number N of supernovae will be measured. Each measurement provides an estimate of the luminosity distance $\bar{D}_L(z)$, but the estimate is dominated by the intrinsic scatter due to the variation of the absolute luminosity. The measurement uncertainty on the luminosity distance can be beaten down by N-independent measurements of supernovae at the same redshift, such that we expect to measure the background luminosity distance $\bar{D}_L(z)$ precisely, provided that the number N of supernovae is sufficiently large ($N \to \infty$).

However, this standard picture is *incorrect*: What we measure is not the background luminosity distance $\bar{D}_L(z)$, but the number weighted luminosity distance $\mathcal{D}(\mathbf{x})$, where our observable $D(\mathbf{x})$ in Eq. (2.17) corresponds to the luminosity distance including the perturbation and the observed data $\mathcal{D}_i^{\rm obs}$ in Eq. (2.15) is the set of luminosity distances weighted by the host galaxy number count, as described by $\mathcal{D}(\mathbf{x})$ in Eqs. (2.18) and (2.19). We emphasize again that the supernova measurements include not only the background luminosity distance, but also the perturbations in Eq. (2.21), which represent the fluctuations in the luminosity distance and the host galaxy. Because of this, even in this idealized situation, where we can beat down the intrinsic scatter completely with infinite number of supernova observations, there exists a cosmic variance limit to the luminosity distance measurements. Therefore, it is important to compute this cosmic variance limit (or the maximum cosmological information contents) derivable from supernova observations in a single redshift bin.[3]

This maximum cosmological information content is derived in Eq. (4.11) under the assumption that the underlying fluctuations are Gaussian and linear. Focusing on the background part, we notice that the variance on the background measurements is the monopole power spectrum C_0:

$$F_{\mu\nu} \propto \frac{4\pi}{C_0}\left(\frac{\partial \ln \bar{D}(z)}{\partial p_\mu}\right)\left(\frac{\partial \ln \bar{D}(z)}{\partial p_\nu}\right) , \tag{5.1}$$

as in Eq. (4.8). Observers derive the best estimate of the background luminosity distance by averaging over the sky coverage in Eq. (2.16), and this estimate is represented by $\langle \mathcal{D} \rangle_\Omega (z)$ in Eq. (2.23), which includes the background luminosity distance and the monopole perturbation $\delta\mathcal{D}_0(z)$ in Eq. (2.24). The monopole is correlated

$$\langle \delta\mathcal{D}_0 \delta\mathcal{D}_0 \rangle \equiv \frac{\left\langle |a_{00}|^2 \right\rangle}{4\pi} = \frac{C_0}{4\pi} , \tag{5.2}$$

and since we only have access to one light cone, the monopole power spectrum sets the cosmic variance limit to the background measurements.

Once we account for the fact that our luminosity distance measurements include perturbations, it is inevitable that the measurements are limited by cosmic variance. The standard way of estimating cosmic variance is to compute

$$\sigma^2_{\rm std} := \langle \delta\mathcal{D}(\mathbf{x}) \delta\mathcal{D}(\mathbf{x}) \rangle \equiv \xi(0) , \tag{5.3}$$

which we refer to as the standard variance. However, note that the fluctuations of the host galaxies are often ignored in literature. Using the angular decomposition in Eq. (3.13), we derive

$$\sigma^2_{\rm std} = \xi(0) = \sum_l \frac{2l+1}{4\pi} C_l \geq \frac{C_0}{4\pi} , \tag{5.4}$$

[3] In [59, 60], similar arguments are presented, regarding the cosmic variance limit. They computed the cosmic variance on the angle average $\langle \mathcal{D}_L \rangle_\Omega (z)$ of the luminosity distances, while ignoring the linear-order monopole contribution and the host galaxy fluctuation, but focusing on the second-order relativistic contributions. So, the cosmic variance obtained in [59, 60] is smaller than our estimate in this section.

where we used $P_l(1) = 1$. Given the correct formula for $\delta\mathcal{D}$, the standard variance is indeed larger than the real cosmic variance limit $C_0/4\pi$, because $\sigma^2_{\rm std}$ is the variance of the luminosity distance fluctuations at each spatial point, while the variance we need for the background estimate is the variance on the angle-averaged fluctuations. We can gain further insight by computing the cosmic variance in configuration space

$$\langle \delta\mathcal{D}_0 \delta\mathcal{D}_0 \rangle := \xi_0 = \int \frac{d\Omega_{12}}{4\pi}\, \xi(\hat{\mathbf{n}}_{12}) = \frac{1}{2}\int_0^\pi d\theta\ \sin\theta\, \xi\left(r = 2\bar{r}_z \sin\frac{\theta}{2}; z\right) \leq \xi(0)\,. \tag{5.5}$$

The monopole correlation $\xi_0(z)$ is literally the angle average of the full three-dimensional correlation function $\xi(r;z)$ on the light cone, and we note that the equality holds only at the tip of the light cone ($z = 0$).

Without detailed calculations, we can obtain a rough estimate of the cosmic variance limit. In Eq. (2.21), the perturbation $\delta\mathcal{D}(\mathbf{x})$ in the observed data set is composed of the fluctuation δD in the luminosity distance and the fluctuation δ_g in the host galaxy clustering. The fluctuation in the luminosity distance was computed, properly accounting for the relativistic contributions (see, e.g., [37–42]), where it was shown that the velocity contributions are larger than the gravitational potential contribution. So, it is clear that the dominant contribution to the supernova observations comes from the fluctuation in the host galaxy clustering, which is in proportion to the matter density fluctuation. However, this density contribution to the monopole $\delta\mathcal{D}_0$ is cancelled by δN, so that the leading contribution to the cosmic variance is the velocity. Considering a simple estimate $\sigma^2_v \simeq 10^{-5}$—$10^{-4}$ at low redshift, a percent level cosmic variance is expected, while at high redshift the lensing contribution is larger than the velocity contribution (see also [59–61], where similar results have been obtained).

Supernova observations, however, are not limited to a single redshift bin, but in general cover a range of redshifts. Furthermore, we are somewhat less interested in the background luminosity distances $\bar{D}_L(z)$ at each redshift, but more in deriving cosmological parameter constraints from the supernova measurements over the redshift ranges. One must therefore consider the full $\mathcal{D}_L(z)$ function and the full light-cone Fisher analysis of section 4.3. As emphasized, even with an infinite number of supernova observations, we cannot perfectly measure the background luminosity distances over a range of redshifts, and there *exist* a maximum cosmological information content set by the cosmic variance limit. To put it differently, there exists a minimum error for cosmological parameter estimation from supernova observations as we have only one light cone at our disposition (up to a maximum redshift). This question will be investigated in detail in future work [62].

5.2 Cosmic microwave background anisotropies: Do we know the background CMB temperature?

In observations of cosmic microwave background anisotropies, we measure the CMB temperature $T(\hat{\mathbf{n}})$ as a function of angular position $\hat{\mathbf{n}}$ in the sky. As discussed in section 2.3, our observable $D(\hat{\mathbf{n}})$ in Eq. (2.17) corresponds to the CMB temperature $T(\hat{\mathbf{n}})$, and the source fluctuation δ_g is absent, as it is unbiased or we explicitly solve the temperature evolution using the Boltzmann equation. Moreover, we need to pay attention to the fact that the temperature measurements give $T(\hat{\mathbf{n}}) = \bar{T}[1 + \Theta(\hat{\mathbf{n}})]$, or the sum of the background temperature $\bar{T}$ and its fluctuation $\Theta(\hat{\mathbf{n}})$. Therefore, the full cosmological information contents are described by Eq. (4.11)

$$F_{\mu\nu} = \frac{4\pi}{C_0}\left(\frac{\partial \ln \bar{T}}{\partial p_\mu}\right)\left(\frac{\partial \ln \bar{T}}{\partial p_\nu}\right) + \sum_{l=0}^{\infty} \frac{2l+1}{2}\frac{\partial}{\partial p_\mu}\left(\ln \bar{T}^2 C_l\right)\frac{\partial}{\partial p_\nu}\left(\ln \bar{T}^2 C_l\right), \tag{5.6}$$

to contrast to the standard Fisher information matrix for CMB

$$F^{\rm std}_{\mu\nu}=\sum_{l=2}^{\infty}\frac{2l+1}{2}\frac{\partial}{\partial p_\mu}\bigg(\ln C_l\bigg)\frac{\partial}{\partial p_\nu}\bigg(\ln C_l\bigg)~, \tag{5.7}$$

where we sum the standard Fisher matrix from the quadrupole $l=2$. The standard Fisher information matrix can be derived from the full Fisher information, if we assume that the background CMB temperature $\bar T$ is precisely known and drops out of the model parameters, and if we assume that the monopole and the dipole contain *no* cosmological information. However, as we argue, neither of these assumptions is correct.

In CMB observations, we obtain the "background" CMB temperature $\langle T\rangle^{\rm obs}$ by averaging $T(\hat{\mathbf{n}})$ over the sky as defined in Eq. (2.28). As emphasized in Eq. (2.24), however, the monopole fluctuation Θ_0 is not zero (but note $\Theta_0^{\rm obs}=0$), and hence the observed CMB temperature $\langle T\rangle^{\rm obs}$ we use is *not* the background CMB temperature $\bar T$ (or $\bar\rho_\gamma$). With the Fisher information matrix $F_{\mu\nu}$ in Eq. (5.6), our estimate of $\bar T$ is subject to the cosmic variance set by the monopole C_0 of the power spectrum, and this part is ignored in the standard analysis $F^{\rm std}_{\mu\nu}$. We suspect that the monopole is ignored, because it is already absorbed in $\langle T\rangle^{\rm obs}$. The observed dipole Θ_1 also contains cosmological information, as it is a measure of the relative velocity between the observer and the CMB fluid, all of which can be predicted in a given cosmological model. But the dipole, being mainly due to our local velocity, is subject to significant nonlinear clustering and galaxy formation which we cannot compute in detail.

We proceed to qualitatively compute the impact on the standard analysis, ignoring the monopole and the dipole contributions in $F_{\mu\nu}$. Given N_p-number of cosmological parameters, we need to consider one extra parameter or the background CMB temperature $\bar T$, such that the full parameter analysis contains N_p+1 parameters with $p_0:=\ln\bar T$ and the full Fisher information matrix is given as

$$F_{\mu\nu}=\begin{pmatrix}F_{00} & F_{0\sigma}\\ F_{\rho0} & F_{\rho\sigma}\end{pmatrix}~,\qquad\qquad \mu,\nu\in(0,1,\cdots,N_p)~,\qquad \rho,\sigma\in(1,\cdots,N_p)~, \tag{5.8}$$

where $F_{\rho\sigma}\equiv F^{\rm std}_{\rho\sigma}$ and the extra components in addition to the standard Fisher information matrix are

$$F_{00}=\frac{4\pi}{C_0}+\sum_l\frac{2l+1}{2}\left(2+\frac{\partial\ln C_l}{\partial\ln\bar T}\right)^2~, \tag{5.9}$$

$$F_{\rho0}=\sum_l\frac{2l+1}{2}\left(\frac{\partial}{\partial p_\rho}\ln C_l\right)\left(2+\frac{\partial\ln C_l}{\partial\ln\bar T}\right)~. \tag{5.10}$$

While the power spectrum C_l decays exponentially at high l and the summation over l with the $\bar T$-derivative converges, these extra components are expected to be large. To derive the information loss in the standard CMB analysis, we need to marginalize over the background CMB temperature. The full covariance matrix or the inverse of the full Fisher matrix is

$$F^{-1}_{\mu\nu}=\begin{bmatrix}\left(F_{00}-F_{0\epsilon}F^{-1}_{\epsilon\delta}F_{\delta0}\right)^{-1} & -\frac{F_{0\epsilon}}{F_{00}}\left(F_{\epsilon\sigma}-F_{\epsilon0}F_{0\sigma}/F_{00}\right)^{-1}\\ -\frac{F^{-1}_{\rho\kappa}F_{\kappa0}}{F_{00}-F_{0\epsilon}F^{-1}_{\epsilon\delta}F_{\delta0}} & \left(F_{\rho\sigma}-F_{\rho0}F_{0\sigma}/F_{00}\right)^{-1}\end{bmatrix}~,\qquad \epsilon,\delta,\kappa\in(1,\cdots,N_p)~, \tag{5.11}$$

and after marginalizing over the background CMB temperature $\bar T$ (or p_0) we invert the marginalized covariance matrix to obtain the reduced Fisher information matrix for N_p-parameters as

$$\tilde F_{\rho\sigma}=F^{\rm std}_{\rho\sigma}-\frac{F_{\rho0}F_{0\sigma}}{F_{00}}~. \tag{5.12}$$

It is clear that some information is lost, compared to the standard Fisher matrix $F^{\rm std}_{\rho\sigma}$. However, the detailed analysis of the information loss and its impact on cosmological parameters will require extensive numerical investigations based on the Boltzmann equation solvers, which is beyond the scope of the present paper. We defer this investigation for future work [63].

5.3 3D Weak gravitational lensing and tomography

3D weak lensing was developed [55] for the first time to utilize the additional redshift information in lensing surveys. Traditionally, shape measurements are made for individual galaxies. While their angular positions are immediately available, the redshift is often unavailable, and the theoretical predictions are then projected along the line-of-sight as in section 2.4 to compare to the observations (see [64–67] for review). However, even in this case, we need the average radial distribution $d\bar{N}/dzd\Omega$ of the source galaxies to correctly compute the theoretical expectation $\mathcal{D}^{\rm pro}$ in Eq. (2.33), and this is often achieved in observations with spectroscopic or photometric redshift measurements for a subset of the source galaxies. Furthermore, the recent technological advances allow fairly accurate photometric redshift measurements for individual galaxies in lensing surveys (see, e.g., [8, 9, 68, 69]), and this additional radial information is indeed useful to handle the systematic errors in lensing data such as the intrinsic alignments [70, 71].

Using this extra information in the radial distribution, a lensing tomography was proposed [72] and is now widely used in lensing surveys (see, e.g., [69]), in which the source galaxies are grouped into several radial bins according to their redshifts and weak lensing measurements are made for individual bins to obtain their auto and cross correlations among the radial bins. Compared to the traditional weak lensing case, the covariance can be readily extended, and the orthonormality condition in the tomographic lensing is then

$$(\mathbf{C}\,\mathbf{K})_{12} = \delta_{z_1 z_2}\delta^D(\Omega_1 - \Omega_2)d\Omega_1 \,, \tag{5.13}$$

where z_i, z_j indicate the tomographic bins. There exists an extra structure for radial bins in the orthonormal relation of tomographic lensing, and this is to be contrasted to Eq. (3.10) for the traditional weak lensing. The maximum cosmological information contents can be quantified by the Fisher matrix, which takes the same form as in the traditional weak lensing, but with the angular power spectrum $C^{\rm std}_l$ now replaced with $C^{\rm tomo}_l$ and the inverse angular power spectrum $\tilde{C}^{\rm std}_l$ with $\tilde{C}^{\rm tomo}_l$

$$C^{\rm tomo}_l := \begin{pmatrix} C^{11}_l & C^{12}_l \\ C^{12}_l & C^{22}_l \end{pmatrix} \,, \qquad \tilde{C}^{\rm tomo}_l := \left(C^{\rm tomo}_l\right)^{-1} \,, \tag{5.14}$$

where we assumed there are two tomographic bins. This recovers the original formulation in [72].

3D weak lensing can be viewed as the tomographic weak lensing in the limit, wheres the number of tomographic bins becomes infinite. In this sense, 3D weak lensing [55] provides the most comprehensive method to use the full information available in lensing surveys. It decomposes the angular position of the shape measurements on the sky in terms of spherical harmonics and the radial position in terms of spherical Bessel function, as described in section 3.3.2. Since more information is used in 3D weak lensing than in the traditional method or tomographic method, more cosmological information can be extracted in 3D weak lensing, and it is important to quantify the net increase in the cosmological information. For illustration, the cosmological information in 3D weak lensing was approximated [55] as the two-dimensional one in Eq. (4.12) for each Fourier mode to be summed over, and it was found that $\sim 30\%$ improvements can be achieved in measuring the underlying matter density power spectrum, though the number depends on the characteristics of the survey. However, this procedure of computing the information contents essentially ignores the radial correlation between the observed data points, and the correct cosmological information in 3D weak lensing needs to be computed by using the full three-dimensional Fisher information matrix in Eq. (4.23), instead of Eq. (4.12). The detailed analysis of cosmological information in 3D weak lensing and tomography will be investigated in future work [73].

5.4 Cosmic variance on the baryon density $\bar{\rho}_b$: Missing baryons in the local Universe

Recent measurements of the cosmic microwave background anisotropies and the light element abundance from the big bang nucleosynthesis yield a very precise value for the background baryon density $\bar{\rho}_b$ today. While the high-redshift measurements of the Lyman alpha forests yields a baryon density consistent with the background baryon density $\bar{\rho}_b(z)$ at the corresponding redshift, it is well-known that the baryon density in the local Universe or low redshift accounts for only about a half of $\bar{\rho}_b$, and this issue is known as the missing baryon problem [74] (see [75] for a recent review). The baryons at low redshift are expected to be in the warm-hot intergalactic medium (WHIM), and they are notoriously difficult to observe in optical or X-ray telescopes (see [76, 77], however, recent measurements of O_{VII} in soft X-rays). In light of the formalism developed in this work, we are interested in estimating the cosmic variance on the observations of the baryon density at low redshift, rather than proposing another solution to the missing baryon problem. At low redshift, the light cone volume is small and observations are subject to considerable cosmic variance. This also applies to the measurements of the baryon density, as observations of the baryon density $\rho_b(\mathbf{x}) = \bar{\rho}_b(z)(1+\delta_b)$ include not only the background baryon density $\bar{\rho}_b(z)$, but also its perturbation δ_b.

Consider estimating the baryon density by observing gas clouds (or WHIM), in which the direct observables are often surrogates for the baryon density such as the oxygen number density $D = n_{O_{VII}}$ (in a specific ionization state) and the host galaxy fluctuation in this case becomes the gas cloud fluctuation $\delta_g = \delta_{\rm WHIM}$. The oxygen abundance needs to be converted to the overall baryon density, but here we simplify the situation by assuming that this procedure is straightforward and we use $D = \rho_b(\mathbf{x})$. Since our interest is in one number or the (background) baryon density $\bar{\rho}_b$ at $z = 0$, we can scale out the redshift dependence, such that we obtain the observational estimate of $\bar{\rho}_b^{\rm obs}$ by averaging all the measurements over the light cone volume in similarity to Eq. (2.16):

$$\bar{\rho}_b^{\rm obs} \equiv \bar{\mathcal{D}}^{\rm obs} := \frac{1}{N}\sum_{i=1}^{N} \mathcal{D}_i^{\rm obs}(1+z_i)^3 \,, \tag{5.15}$$

where N is indeed a few for the case of local baryon density measurements.

This average can be modeled as in Eqs. (2.20) and (2.21) by using $\bar{D}(z) = (1+z)^3\bar{\rho}_b(z)$ and $\delta_g = \delta_{\rm WHIM}$. While the full information content in these observations is described by Eq. (4.23), we focus on the dominant contribution or the cosmic variance on the background baryon density $\bar{\rho}_b$ in Eq. (4.15). Assuming that the cosmological parameters other than $\bar{\rho}_b$ are known, and the observations are at low redshift, we first compute the Fourier kernel in Eq. (4.16)

$$\mathcal{G}_\mu = \int dz\; j_0(k\bar{r})\frac{\partial}{\partial\bar{\rho}_b}\ln\hat{\bar{\mathcal{D}}}(z) = \frac{1}{\bar{\rho}_b}\int dz\; j_0(k\bar{r}) \simeq \frac{\Delta z}{\bar{\rho}_b}j_0\left(\frac{k\Delta z}{H_0}\right) \,, \tag{5.16}$$

and the Fisher information matrix in Eq. (4.15) becomes

$$F_{\mu\nu} \approx \left(\frac{4\pi\Delta z}{\bar{\rho}_b}\right)^2 \int dk \int dk'\, \frac{kk'}{2\pi^2}\tilde{S}_0(k,k')j_0\left(\frac{k\Delta z}{H_0}\right)j_0\left(\frac{k'\Delta z}{H_0}\right) \,, \tag{5.17}$$

where Δz is the survey depth in redshift. Due to the rapid oscillations and decay of the spherical Bessel function, the integral receives appreciable contributions only over the k-range, where the argument of the spherical Bessel function is less than about 2,

$$0 \leq \frac{k\Delta z}{H_0} \leq 2\,. \tag{5.18}$$

Therefore, the Fisher matrix is

$$F_{\mu\nu} \approx \frac{16H_0^4}{\Delta z^2 \bar{\rho}_b^2} \tilde{S}_0(\bar{k}, \bar{k})\,, \tag{5.19}$$

where we used $j_0^2 \simeq 0.5$ over the range and $\Delta k \simeq 2H_0/\Delta z$ and defined the representative wave number $\bar{k} := H_0/\Delta z$.

To proceed, we make a series of assumptions to compute the inverse spherical power spectrum $\tilde{S}_0(k, k')$ by using Eq. (3.30). As in section 5.1, the gas cloud (or WHIM) is dominated by the matter density clustering $\delta_g \simeq \delta_m$, and the spherical power spectrum of the matter density fluctuation at low redshift is isotropic $S_l(k, k') \approx \delta^D(k - k')P_m(k)$ (see Appendix B). The Fourier angular kernel in Eq. (3.25) is non-vanishing

$$\mathcal{F}_l(k_1, k_2) := \int dz\, j_l(k_1\bar{r})j_l(k_2\bar{r}) \approx \frac{1}{2}\Delta z\,, \tag{5.20}$$

only over $0 \leq k_1, k_2 \leq \Delta k$, and the integral equation (3.30) becomes

$$1 \approx \frac{8}{\pi^2}\frac{H_0^7}{\Delta z^5}\tilde{S}_0(\bar{k}, \bar{k})P_m(\bar{k}) \simeq \frac{16H_0^4}{\Delta z^2}\tilde{S}_0(\bar{k}, \bar{k})\sigma_R^2\,, \tag{5.21}$$

where we defined the rms matter fluctuation smoothed by a top-hat radius $R = \Delta z/H_0$:

$$\sigma_R^2 := \int d\ln k\, \frac{k^3}{2\pi^2}P_m(k)j_0^2(kR) \simeq \left(\frac{H_0}{\Delta z}\right)^3 \frac{P_m(\bar{k})}{2\pi^2}\,. \tag{5.22}$$

The full Fisher matrix simplifies accordingly, and the uncertainty on the baryon density becomes

$$F_{\mu\nu}^{-1/2} \approx \bar{\rho}_b\sigma_R\,. \tag{5.23}$$

The cosmic variance is driven by the matter density fluctuations, and at low redshift it is similar to the rms fluctuation smoothed with the scale set by the redshift depth. With $\Delta z \simeq 0.1$, the comoving radius is about $R \simeq 300\ h^{-1}\mathrm{Mpc}$, and the rms fluctuation $\sigma_R \simeq 0.06$. The rms fluctuation further decreases to $\sigma_R \simeq 0.002$, as the survey depth increases to $\Delta z \simeq 0.3$.

Given the measurement uncertainties and the amount of missing baryons in the local Universe, the cosmic variance contributes only a small fraction to the problem. However, it is important to know that the missing 50% correspond to about $8.3\sigma_R$. Furthermore, our simple estimate is based on many simplifying assumptions: First, we computed the maximum cosmological information in an idealized survey, where an infinite number of measurements can be made. However, real observations of the local baryon density are indeed based on a few sight lines towards bright background sources, dramatically increasing the sample variance in real observations, compared to our cosmic variance limit. Second, the location of the observers is not a random place in the Universe, but a highly biased placed (or a Milky-way sized halo). Furthermore, while we used the linear theory to compute the cosmic variance, the real analysis has to account for the nonlinear effects of galaxy clustering. These two effects will greatly increase the variance in the local baryon density measurements. The detailed analysis of all these effects will require numerical simulations, and it will be investigated in future work [78].

5.5 Galaxy power spectrum in a cube vs spherical power spectrum on a light cone

The standard galaxy power spectrum analysis proceeds as though the survey volume is in the hypersurface of simultaneity and a Fourier decomposition is made in the rectangular box. The uncertainties of the power spectrum estimates are further reduced by the number of Fourier modes available in the survey volume. This is qualitatively correct, if the survey volume is small enough, but it becomes inaccurate as we look into the large scale modes and the survey volume becomes larger. Here we make the connection of this traditional power spectrum analysis to our spherical power spectrum analysis, by which we quantify what conditions are needed to justify the small volume.

We first need to compute the inverse spherical power spectrum $\tilde{S}_l(k,k')$ given in Eq. (3.30). Under the assumption that the sky coverage is small and the redshift depth is shallow, we will make use of a series of manipulations based on the rapid oscillating properties of the spherical Bessel function, called the Limber approximation [79, 80]

$$\int_0^\infty dx\, x^2 f(x) j_l(\alpha x) j_l(\beta x) \simeq \frac{\pi}{2\alpha^2}\delta^D(\alpha-\beta) f\left(x = l+\frac{1}{2}\right) , \tag{5.24}$$

where the function $f(x)$ is assumed to be slowly varying over the range relevant to the integration and it becomes the identity when $f(x)$ is a constant. Assuming also $S_l(k,k') \approx \delta^D(k-k')P_m(k)$ and performing the integration over k_1, we obtain

$$\left(\frac{2}{\pi}\right)\int dk_1'\; k_1' k_B \tilde{S}_l(k_1',k_A) \times \int dr\, \frac{H^2}{\bar{r}^2} j_l(k_1'\bar{r}) j_l(k_B\bar{r}) P(k_\star) = \delta^D(k_A - k_B) , \tag{5.25}$$

where the Hubble parameter was introduced by converting the integration variable from dz to $d\bar{r}$ and the star indicates that the Fourier mode is evaluated under the condition $k_\star\bar{r} = l+1/2$. Further assuming that the inverse spherical power spectrum is $\tilde{S}_l(k,k') \approx \delta^D(k-k')\tilde{S}_l(k)$, and integrating over k_A, the closed equation can be manipulated as

$$1 \simeq \int d\bar{r}\, \frac{H^2}{\bar{r}^2} P(k_\star) \times \frac{2}{\pi}\int dk_A\; k_A^2 \tilde{S}_l(k_A) j_l(k_A\bar{r}) j_l(k_A\bar{r}) . \tag{5.26}$$

Applying the same trick for the spherical Bessel function one more time to the integration over $\bar{r}$ and simplifying the remaining integral with the Dirac delta function, we derive the inverse spherical power spectrum

$$\tilde{S}_l(k) \simeq \left(\frac{\bar{r}^2}{H}\right)^2_\star P^{-1}(k) , \tag{5.27}$$

where the star indicates now that the comoving radius is evaluated under the condition $k\bar{r}_\star = l+1/2$. For a small survey volume, where the flat-sky approximation is accurate and the redshift evolution is negligible, the inverse spherical power spectrum is literally the inverse of the power spectrum, but with the volume factor to compensate for the dimensionful quantity.

Having derived the inverse spherical power spectrum, we are now in a position to tackle the more complicated equation for the full Fisher information matrix in Eq. (4.23). Under the same assumption that the survey volume is small, we can ignore the Fourier angular kernel

$$\mathcal{N}_{l,\mu}(k_1,k_2) = \int dz\; j_l(k_1\bar{r}) \frac{\partial}{\partial p_\mu} j_l(k_2\bar{r}) \simeq 0 . \tag{5.28}$$

In addition, we assume that the background galaxy number density $\hat{\bar{\mathcal{D}}} := \bar{n}(z)$ is known, so that we can ignore the other Fourier angular kernel

$$\mathcal{H}_{l,\mu}(k_1,k_2) = \int dz\; j_l(k_1\bar{r}) j_l(k_2\bar{r}) \frac{\partial}{\partial p_\mu} \ln\bar{n}(z) \simeq 0 , \tag{5.29}$$

though we can only measure $\bar{n}(z)$ up to the monopole contribution. This assumption also eliminates the mean contribution to the Fisher matrix, and only the covariance of the galaxy number counts contributes to the Fisher matrix. Furthermore, we only consider the monopole power spectrum $S_0(k,k')$ to make a connection to the angle-averaged power spectrum. The traditional power spectrum analysis proceeds as if the (small) survey volume is embedded in a cubic volume of hypersurface with the origin at the center of the cubic volume, while the observer is indeed at a distance, so that the line-of-sight direction is considered fixed over the survey volume. Under this assumption, the observed power spectrum is well approximated as the redshift-space power spectrum described by the Kaiser formula [25], and the redshift-space power spectrum has the monopole, the quadrupole and the hexadecapole only. The information contents for this monopole power spectrum were derived [5, 6]:

$$F_{\mu\nu}^{\rm std} = 2\pi \int d\ln k \left(\frac{k}{2\pi}\right)^3 V_{\rm eff} \left[\frac{\partial}{\partial p_\mu} \ln P(k)\right] \left[\frac{\partial}{\partial p_\nu} \ln P(k)\right] , \tag{5.30}$$

where $P(k)$ represents the monopole power spectrum and

$$V_{\rm eff} := \int d^3x \left[\frac{\bar{n}(x)P(k)}{1+\bar{n}(x)P(k)}\right]^2 \tag{5.31}$$

is the survey volume in our idealized case ($\bar{n} \to \infty$).

This monopole power spectrum or the angle-average of the power spectrum in Fourier space can be considered as the observed angle-average of the power spectrum if the observer is located at the center of the survey volume. Under this assumption, the monopole power spectrum corresponds to our monopole spherical power spectrum $S_0(k,k')$. Therefore, these assumptions greatly simplify the Fisher information matrix in Eq. (4.23) to

$$F_{\mu\nu} \simeq \frac{(4\pi)^4}{2} \left(\int d\ln k_1 \frac{k_1^3}{2\pi^2} \cdots \int d\ln k_4 \frac{k_4^3}{2\pi^2}\right) \tilde{S}_0(k_1)\tilde{S}_0(k_3)\mathcal{F}_0^{12}\mathcal{F}_0^{23}\mathcal{F}_0^{34}\mathcal{F}_0^{41} \frac{\partial}{\partial p_\mu}P(k_2)\frac{\partial}{\partial p_\nu}P(k_4) . \tag{5.32}$$

Our strategy is again to apply the Limber approximation Eq. (5.24) multiple times to simplify the integration over the wavevector and redshift. We first perform the integration over k_1 and k_3 with Eq. (5.24) and simplify the Dirac delta function to derive

$$\begin{aligned} F_{\mu\nu} \simeq\ & \frac{(4\pi)^2}{2} \int d\ln k_2 \frac{k_2^3}{2\pi^2} \int d\ln k_4 \frac{k_4^3}{2\pi^2} \int d\bar{r}_i \left(\frac{H_i}{\bar{r}_i}\right)^2 \int d\bar{r}_j \left(\frac{H_j}{\bar{r}_j}\right)^2 \\ & \times j_0(k_2\bar{r}_i)j_0(k_2\bar{r}_j)j_0(k_4\bar{r}_j)j_0(k_4\bar{r}_i)\tilde{S}_0\left(k_1=\frac{1}{2\bar{r}_i}\right)\tilde{S}_0\left(k_3=\frac{1}{2\bar{r}_j}\right)\frac{\partial}{\partial p_\mu}P(k_2)\frac{\partial}{\partial p_\nu}P(k_4) . \end{aligned} \tag{5.33}$$

Applying Eq. (5.24) to the integration over k_4 and using the expression for the inverse spherical power spectrum $\tilde{S}_0(k)$ in Eq. (5.27), we obtain

$$F_{\mu\nu} \simeq 2\pi \int d\ln k \frac{k^3}{2\pi^2}\frac{\partial}{\partial p_\mu}P(k) \int d\bar{r}\, \bar{r}^2 j_0(k\bar{r})j_0(k\bar{r})P^{-1}(k_\star)\frac{\partial}{\partial p_\nu}\ln P(k_\star) , \tag{5.34}$$

where $k_\star := 1/2\bar{r}$. The radial integral is then re-arranged by using Eq. (5.24) to arrive at the desired equation of the standard power spectrum analysis

$$F_{\mu\nu} \simeq 2\pi \int d\ln k \left(\frac{k}{2\pi}\right)^3 V_{\rm eff} \left[\frac{\partial}{\partial p_\mu} \ln P(k)\right] \left[\frac{\partial}{\partial p_\nu} \ln P(k)\right] , \tag{5.35}$$

where the effective volume is

$$V_{\rm eff} := 4\pi \int d\bar{r}\, \bar{r}^2 j_0^2(k\bar{r}) \simeq \frac{2\pi R}{k^2}\,, \tag{5.36}$$

where R denotes the survey depth and we ignored the lower boundary of the survey. Compared to the spherical Fourier analysis, the standard power spectrum analysis in summary makes a series of approximations: (1) the redshift evolution over the survey volume is negligible, (2) the angular position is constant (distant-observer approximation), (3) the angular diameter distances $\bar{r}$ are independent of cosmological parameters ($\mathcal{N}_{l,\mu} = 0$), (4) the background galaxy number density is known.

It is well known that the redshift-space power spectrum contains more information than just the monopole power spectrum. It is evident now that a lot more cosmological information is available in galaxy surveys and not all the information has been utilized in the traditional power spectrum analysis. The detailed power spectrum analysis will be investigated in future work [81].

6 Discussion and Summary

In this paper, we have developed a theoretical framework to describe cosmological observables on the light cone and we have derived the Fisher information matrix to quantify the maximum cosmological information obtainable from cosmological observables such as the luminosity distance, weak gravitational lensing, galaxy clustering, and the cosmic microwave background (CMB) anisotropies. As all the cosmological observables contain perturbations, their measurements are subject to the cosmic variance, and in computing the cosmic variance, we have taken into account that the survey geometry is the light cone volume. In section 5, we have discussed in detail the impact of our formulation on the cosmological information contents for five different cosmological observables, in comparison to the standard analysis. Our main findings are as follows:

- Our theoretical framework provides a unified description of angular observables, observables with redshift information, and their variants such as the projected observables. Moreover, it accounts for the fact that observables are often obtained with weights given by the number counts of host galaxies. The measurements of type Ia supernovae are, for instance, modulated not only by the fluctuations in the luminosity distance itself, but also by the spatial correlation of the host galaxies, as we can only have supernovae in a host galaxy. While the latter is often ignored in literature, it is indeed the dominant source of perturbations.

- To properly quantify the cosmological information contents that can be derived from a given observable, we have deployed the Fisher information technique and assumed a Gaussian probability distribution. As the cosmological observables can be measured over a range of redshift, we need to account for their three-dimensional correlation on the light cone and to derive its inverse in computing the Fisher information matrix. In the standard picture, where our survey volume is treated as a cubic box, the Fourier analysis provides the best way for this task, and the inverse of the power spectrum is trivial, as each Fourier mode is independent. However, in the real Universe, where the survey volume is on the past light cone, the radial and the angular positions carry different information. To properly accommodate this, we have used the spherical Fourier analysis and derived for the first time the closed equation (3.30) for the inverse of the three-dimensional correlation of the cosmological observables. We have fully taken into account that the lightcone geometry breaks translation invariance in the radial direction and therefore radial Fourier modes are correlated.

- Given the inverse spherical power spectrum, it is straightforward to derive the Fisher information matrix for three-dimensional cosmological observables. To obtain rough estimates for the impact of cosmic variance, we have applied it to supernova observations, local baryon density measurements, 3D weak gravitational lensing, and galaxy clustering, all of which are three-dimensional on the light cone and are correlated (see section 5 for detailed discussions). While the first two are often thought to provide measurements of background quantities such as the background luminosity distance and the global baryon density, they both measure only the sum of the background and the perturbation together, and hence these measurements are also subject to cosmic variance (sections 5.1 and 5.4).

- For three-dimensional cosmological probes such as galaxy clustering and 3D weak lensing, the spherical Fourier analysis is used to analyze the observables on the light cone, and our Fisher matrix analysis in section 5.5 shows that the standard analysis is based on many simplifying assumptions such as the distant observer, the flat-sky, and no radial correlations.

- Regarding angular cosmological observables such as CMB anisotropies and weak lensing observables, the standard formalism correctly describes these angular observables on the light cone, except one subtlety associated with the observed mean in the CMB temperature. The observed mean of the CMB temperature is obtained by averaging the CMB temperature on the sky, which includes not only the background $\bar{T}$ (an input cosmological parameter), but also the monopole perturbation (a prediction of the model). While the cosmic variance in the observed temperature is expected to be 10^{-5}, its present measurement error [82, 83] is about 2.1×10^{-4} which is similar to the error in e.g. the angular scale subtended by the acoustic peaks [84] which is $\Delta\theta_* = 3 \times 10^{-4}$. This error propagates to the power spectrum measurements into cosmological parameter estimation. A proper analysis of this subtlety in CMB observations will quantify its impact on the cosmological parameter analysis [63].

In this paper, we derived the maximum cosmological information contents from a cosmological observable. This cosmic variance limit arises, because we have a single light-cone volume at our disposition for observations. A way to make maximal use of it is simply to measure the cosmological observables up to higher redshift, at which the light cone volume is large enough to overcome the disadvantage from a single observation point. A more practical solution is the multi-tracer method [85, 86], already developed in the standard analysis, and it can be easily generalized to the light cone analysis. The idea there is to consider several different cosmological observables which trace the same underlying density field, so that by measuring those observables, one can effectively increase the sampling rate and may eliminate the stochasticity completely for certain cosmological parameters in an idealized case.

Acknowledgments

We acknowledge useful discussions with Avi Loeb and Matias Zaldarriaga. We acknowledge support by the Swiss National Science Foundation. J.Y. and E.M. are further supported by a Consolidator Grant of the European Research Council (ERC-2015-CoG grant 680886).

A Spherical Fourier analysis with the observed redshift as a dimensionless radial distance

We have developed the spherical Fourier decomposition of the cosmological observables in section 3.3.2 and computed the Fisher information matrix in a light cone volume in section 4.3. In this formalism, the cosmological observable on a light cone can be naturally decomposed. However, there exists one subtlety, albeit not a problem: The Fourier analysis of the radial modes is based on the comoving distance $\bar{r}$ and its Fourier mode k:

$$\mathbf{x} = \bar{r}_z\hat{\mathbf{n}} , \qquad\qquad k \sim 1/\bar{r}_z , \qquad [k] = L^{-1} , \tag{A.1}$$

and the conversion of the observed redshift z into the comoving distance $\bar{r}_z$ involves a cosmological model. Here we present an alternative method to perform the spherical Fourier analysis, in which we construct a new observer coordinate and its Fourier counterpart:

$$\mathbf{x} = z\hat{\mathbf{n}} , \qquad\qquad k \sim 1/z , \qquad [k] = 1 . \tag{A.2}$$

This is rather unconventional, but we can readily apprehend its advantage: The spherical power spectrum can be constructed out of the raw observed data in a model independent way. A slight disadvantage arises when we compare to the standard theoretical predictions. For example, we understand well how much power is at $k = 1h\mathrm{Mpc}^{-1}$, while we will need a model-dependent conversion to understand the power at the dimensionless scale $k = 1$. A simple calculation shows that $k = 1\,h\mathrm{Mpc}^{-1}$ would correspond to $\bar{r} \approx 2\pi\ h^{-1}\mathrm{Mpc}$, which would then correspond to the redshift $z \approx 0.002$ in a ΛCDM with $\Omega_m = 0.3$, hence the dimensionless Fourier number would be $k \approx 2\pi/0.002 \simeq 3100$. Though this change amounts to a simple conversion of units at low redshift, the relation between the comoving distance and the redshift is highly nonlinear at high redshift, and the comparison to the standard analysis becomes non-trivial.

The spherical Fourier analysis proceeds almost exactly the same way in section 3.3.2 with the comoving distance $\bar{r}$ replaced by the redshift z. The cosmological observables are now decomposed as

$$\delta\mathcal{D}(\mathbf{x}) := \sum_{lm}\int_0^\infty dk\sqrt{\frac{2}{\pi}}\, kj_l(kz)Y_{lm}(\hat{\mathbf{n}})s_{lm}(k) , \tag{A.3}$$

$$s_{lm}(k) \equiv \int d\Omega\int dz\ z^2\sqrt{\frac{2}{\pi}}\, kj_l(kz)Y^*_{lm}(\hat{\mathbf{n}})\delta\mathcal{D}(\mathbf{x}) . \tag{A.4}$$

The Fourier mode $s_{lm}(k)$ and its spherical power spectrum $S_l(k,k')$ are now dimensionless. For the two-point ξ_{12} and its inverse ζ_{12} correlation functions, the decomposition remains almost unchanged as in Eqs. (3.22) and (3.23)

$$\xi_{12} = 4\pi\sum_{lm}\int dk\int dk'\ \frac{kk'}{2\pi^2}S_l(k,k')j_l(kz_1)j_l(k'z_2)Y_{lm}(\hat{\mathbf{n}}_1)Y^*_{lm}(\hat{\mathbf{n}}_2) , \tag{A.5}$$

$$\zeta_{12} =: 4\pi\sum_{lm}\int dk\int dk'\ \frac{kk'}{2\pi^2}\tilde{S}_l(k,k')j_l(kz_1)j_l(k'z_2)Y_{lm}(\hat{\mathbf{n}}_1)Y^*_{lm}(\hat{\mathbf{n}}_2) , \tag{A.6}$$

which also defines the "inverse" spherical power spectrum $\tilde{S}_l(k,k')$. The relation between the spherical power spectrum and its inverse is exactly the same as in Eq. (3.30), though the Fourier angular kernel is slightly different

$$\mathcal{F}_l(k_1,k_2) := \int dz\ j_l(k_1z)j_l(k_2z) = \frac{\pi}{2(2l+1)}\frac{(k_<)^l}{(k_>)^{l+1}} , \tag{A.7}$$

yet analytically solvable, if the integration is performed from zero to infinity, where $k_>$ is the maximum of k_1 and k_2. Despite this analytic solution, the closed equation for the inverse spherical power spectrum cannot be further simplified, as the spherical power spectrum is an unknown input function. The inverse covariance in the light cone volume is

$$\mathbf{K}_{12} = \frac{4\pi}{\hat{\bar{\mathcal{D}}}(z_1)\hat{\bar{\mathcal{D}}}(z_2)} \sum_l \frac{2l+1}{4\pi} P_l(\gamma_{12}) \int dk \int dk' \, \frac{kk'}{2\pi^2} \tilde{S}_l(k,k') j_l(kz_1) j_l(k'z_2) \, . \tag{A.8}$$

We now proceed to compute the Fisher information matrix in a light cone volume in section 4.3. The contribution of the mean to the Fisher information matrix is exactly the same as

$$\frac{1}{2}\mathrm{Tr}\Big[\mathbf{K}\mathbf{M}_{\mu\nu}\Big] = (4\pi)^2 \int dk \int dk' \, \frac{kk'}{2\pi^2} \tilde{S}_0(k,k') \mathcal{G}_\mu(k) \mathcal{G}_\nu(k') \, , \tag{A.9}$$

with the same Fourier kernel

$$\mathcal{G}_\mu(k) := \int dz \; j_0(kz) \frac{\partial}{\partial p_\mu} \ln \hat{\bar{\mathcal{D}}}(z) \, . \tag{A.10}$$

The variation of the covariance in the Fisher information matrix can be readily obtained simply by replacing $\bar{r}$ with z, and the expression is almost identical to Eq. (4.18), except that the spherical Bessel functions are mow independent of cosmological parameters and they can be pulled out of the derivative with respect to the model parameters. Consequently, the contribution of the covariance to the Fisher information matrix has the same structure with the same Fourier angular kernels:

$$\begin{aligned}\frac{1}{2}\mathrm{Tr}\left[\left(\mathbf{K}\frac{\partial}{\partial p_\mu}\mathbf{C}\right)\left(\mathbf{K}\frac{\partial}{\partial p_\nu}\mathbf{C}\right)\right] &= \left(\frac{2}{\pi}\right)^4 \sum_l \frac{2l+1}{2} \left(\prod_{i=1}^4 \int dk_i \; k_i\right)\left(\prod_{j=1}^4 \int dk'_j \; k'_j\right) \tilde{S}_l(k_1,k_2)\tilde{S}_l(k_3,k_4) \\ &\times \left[\mathcal{F}_l^{21'}\mathcal{H}_{l,\mu}^{32'} S_l(k'_1,k'_2) + \mathcal{F}_l^{32'}\mathcal{H}_{l,\mu}^{21'} S_l(k'_1,k'_2) + \mathcal{F}_l^{21'}\mathcal{F}_l^{32'} \frac{\partial}{\partial p_\mu} S_l(k'_1,k'_2)\right] \\ &\times \left[\mathcal{F}_l^{43'}\mathcal{H}_{l,\nu}^{14'} S_l(k'_3,k'_4) + \mathcal{F}_l^{14'}\mathcal{H}_{l,\nu}^{43'} S_l(k'_3,k'_4) + \mathcal{F}_l^{43'}\mathcal{F}_l^{14'} \frac{\partial}{\partial p_\nu} S_l(k'_3,k'_4)\right] \, .\end{aligned} \tag{A.11}$$

except that one Fourier kernel is identically vanishing:

$$\mathcal{N}_{l,\mu}(k_1,k_2) := \int dz \; j_l(k_1 z) \frac{\partial}{\partial p_\mu} j_l(k_2 z) \equiv 0 \, . \tag{A.12}$$

B Spherical power spectrum on the light cone

Fourier analysis and the power spectrum provide the best way to characterize the initial conditions and their subsequent evolutions at the linear order in perturbations in a hypersurface of simultaneity. However, in observations all the cosmological observables are measured along the past light cone, piercing through different hypersurfaces. Since the Fourier transformation is intrinsically non-local, the Fourier analysis of the cosmological observables, either traditional or spherical, involves non-trivial complications due to the time evolution of perturbations.

To illustrate the point, we assume that our cosmological observable is simply the matter density field $\delta\mathcal{D}(\mathbf{x}) := \delta_m(\mathbf{x})$ located at the observed redshift and angle. The spherical Fourier analysis starts with the decomposition in Eq. (3.18):

$$\delta_m^{\mathrm{obs}}(\mathbf{x}) := \sum_{lm} \int_0^\infty dk \sqrt{\frac{2}{\pi}} \, k j_l(k\bar{r}_z) Y_{lm}(\hat{\mathbf{n}}) s_{lm}^{\mathrm{th}}(k;t_z) \, , \qquad \mathbf{x} = (z, \hat{\mathbf{n}}) \, , \tag{B.1}$$

where we used the superscripts “th” and “obs” to indicate that the left-hand side is obtained in observations and the right-hand side is our theoretical modeling of the observation. The Fourier component $\delta_m^{\rm th}(\boldsymbol{k};t)$ of the matter density, for instance, is a good example of the theoretical quantity, and it is characterized by its power spectrum $P_m^{\rm th}(k;t)$, where the superscripts are often omitted in literature. However, note that these theoretical quantities are computed in a hypersurface, so that the Fourier component of the matter density, for example, implicitly assumes the time-dependence $\delta_m^{\rm th}(\boldsymbol{k};t)$, i.e., Fourier decomposition is performed in a hypersurface of constant t, inaccessible to the observer. Hence, the spherical Fourier component $s_{lm}^{\rm th}(k;t_z)$ has also the time-dependence, where the observed redshift z specifies the hypersurface.

Completely independent of our theoretical modeling, the (observed) spherical Fourier coefficients can be obtained in terms of the (observed) density field $\delta_m^{\rm obs}(\mathbf{x})$ as

$$s_{lm}^{\rm obs}(k) := \int d\Omega \int d\bar{r}\, \bar{r}^2 \sqrt{\frac{2}{\pi}}\, k j_l(k\bar{r}) Y_{lm}^*(\hat{\mathbf{n}}) \delta_m^{\rm obs}(\mathbf{x}) \,, \tag{B.2}$$

where $s_{lm}^{\rm obs}(k)$ is independent of time as the time-dependence is integrated out. Using the spherical Fourier decomposition above, we derive the relation between $s_{lm}^{\rm obs}$ and $s_{lm}^{\rm th}$ as

$$s_{lm}^{\rm obs}(k) = \int dk' \left[\frac{2kk'}{\pi} \int d\bar{r}\, \bar{r}^2 j_l(k\bar{r}) j_l(k'\bar{r})\right] \times s_{lm}^{\rm th}(k';t_z) \,. \tag{B.3}$$

Were it not for the light-cone observation (or the time-dependence) and the finite survey volume, $s_{lm}^{\rm th}$ could be pulled out of the line-of-sight integration, and the relation would simply indicate

$$s_{lm}^{\rm obs}(k) \equiv s_{lm}^{\rm th}(k) \,, \tag{B.4}$$

as desired. Due to the time evolution, however, a non-trivial complication arises for their relation in Eq. (B.3).

Similarly, the (observed) spherical power spectrum can be obtained by considering the ensemble average of the (observed) spherical Fourier components

$$\begin{aligned}\left\langle s_{lm}^{\rm obs}(k) s_{l'm'}^{\rm obs*}(k')\right\rangle = \frac{2}{\pi} \int d\Omega_1 \int d\Omega_2 \int d\bar{r}_1 \int d\bar{r}_2\, \bar{r}_1^2 \bar{r}_2^2\, kk' j_l(k\bar{r}_1) j_{l'}(k'\bar{r}_2) \\ \times Y_{lm}^*(\hat{\mathbf{n}}_1) Y_{l'm'}(\hat{\mathbf{n}}_2) \left\langle \delta_m^{\rm obs}(\mathbf{x}_1) \delta_m^{\rm obs}(\mathbf{x}_2)\right\rangle \,.\end{aligned} \tag{B.5}$$

The (observed) two-point correlation function is then related to the theoretical power spectrum:

$$\left\langle \delta_m^{\rm obs}(\mathbf{x}_1) \delta_m^{\rm obs}(\mathbf{x}_2)\right\rangle = \int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{k}\cdot(\mathbf{x}_1-\mathbf{x}_2)} P_m^{\rm th}(k;z_1,z_2) \,, \tag{B.6}$$

where the power spectrum involves two different hypersurfaces specified by z_1 and z_2 and in linear theory this can be factored out by using the growth factor $D(z)$ normalized at some initial time t_0 as

$$P_m^{\rm th}(k;z_1,z_2) = D(z_1) D(z_2) P_m^{\rm th}(k;t_0) \,. \tag{B.7}$$

Expanding the exponential factor, the ensemble average can be arranged as

$$\begin{aligned}\left\langle s_{lm}^{\rm obs}(k) s_{l'm'}^{\rm obs*}(k')\right\rangle = \delta_{ll'}\delta_{mm'} \int d\tilde{k} \left[\frac{2k\tilde{k}}{\pi} \int d\bar{r}_1\, \bar{r}_1^2 j_l(k\bar{r}_1) j_l(\tilde{k}\bar{r}_1)\right] \\ \times \left[\frac{2k'\tilde{k}}{\pi} \int d\bar{r}_2\, \bar{r}_2^2 j_l(k'\bar{r}_2) j_l(\tilde{k}\bar{r}_2)\right] P_m^{\rm th}(\tilde{k};z_1,z_2) \,,\end{aligned} \tag{B.8}$$

and the (observed) spherical power spectrum is therefore

$$S_l^{\rm obs}(k,k') \equiv \int d\tilde{k}\left[\frac{2k\tilde{k}}{\pi}\int d\bar{r}_1\ \bar{r}_1^2 j_l(k\bar{r}_1)j_l(\tilde{k}\bar{r}_1)\right]\left[\frac{2k'\tilde{k}}{\pi}\int d\bar{r}_2\ \bar{r}_2^2 j_l(k'\bar{r}_2)j_l(\tilde{k}\bar{r}_2)\right] P_m^{\rm th}(\tilde{k};z_1,z_2)\ . \tag{B.9}$$

Again, in the absence of the time-evolution along the light cone in a survey with infinite volume, the power spectrum $P_m^{\rm th}$ could be pulled out of the line-of-sight integrations, and the square brackets simplify to the Dirac delta functions, yielding

$$S_l^{\rm obs}(k,k') = \delta^D(k-k')P_m^{\rm th}(k)\ . \tag{B.10}$$

This simplification is again possible, only if the time-evolution along the light cone is neglected. In linear theory, the (observed) spherical power spectrum can be further simplified by using the growth factor as

$$S_l^{\rm obs}(k,k') = \int d\tilde{k}\ P_m^{\rm th}(\tilde{k};t_0)\mathcal{T}_l(k,\tilde{k})\mathcal{T}_l(k',\tilde{k}) \tag{B.11}$$

where we defined a Fourier kernel

$$\mathcal{T}_l(k,\tilde{k}) := \frac{2k\tilde{k}}{\pi}\int d\bar{r}\ \bar{r}^2 j_l(k\bar{r})j_l(\tilde{k}\bar{r})D(z)\ . \tag{B.12}$$

References

[1] P. J. E. Peebles, *The large-scale structure of the universe*. Princeton University Press, Princeton, 1980.

[2] J. A. Peacock, *Cosmological Physics*. Jan., 1999.

[3] S. Dodelson, *Modern cosmology*. 2003.

[4] R. Durrer, *The Cosmic Microwave Background*. Cambridge University Press, Aug., 2008.

[5] H. A. Feldman, N. Kaiser, and J. A. Peacock, "Power-spectrum analysis of three-dimensional redshift surveys," *Astrophys. J.* **426** (May, 1994) 23–37, arXiv:astro-ph/9304022.

[6] M. Tegmark, "Measuring Cosmological Parameters with Galaxy Surveys," *Phys. Rev. Let.* **79** (Nov., 1997) 3806–3809, arXiv:astro-ph/9706198.

[7] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein, B. Flaugher, K. Honscheid, R. Kron, O. Lahav, P. McDonald, N. Roe, *et al.*, "The DESI Experiment, a whitepaper for Snowmass 2013," *ArXiv e-prints* (Aug., 2013) , arXiv:1308.0847 [astro-ph.CO].

[8] C. W. Stubbs, D. Sweeney, J. A. Tyson, and LSST Collaboration, "An Overview of the Large Synoptic Survey Telescope (LSST) System," in *American Astronomical Society Meeting Abstracts*, vol. 36 of *Bulletin of the American Astronomical Society*, p. 108.02. Dec., 2004.

[9] R. Laureijs, J. Amiaux, S. Arduini, J. . Augueres, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, *et al.*, "Euclid Definition Study Report," *ArXiv e-prints* (Oct., 2011) , arXiv:1110.3193.

[10] J. Green, P. Schechter, C. Baltay, R. Bean, D. Bennett, R. Brown, C. Conselice, M. Donahue, *et al.*, "Wide-Field InfraRed Survey Telescope (WFIRST) Final Report," arXiv:1208.4012.

[11] J. Binney and T. Quinn, "Gaussian random fields in spherical coordinates," *Mon. Not. R. Astron. Soc.* **249** (Apr., 1991) 678–683.

[12] K. B. Fisher, O. Lahav, Y. Hoffman, D. Lynden-Bell, and S. Zaroubi, "Wiener reconstruction of density, velocity and potential fields from all-sky galaxy redshift surveys," *Mon. Not. R. Astron. Soc.* **272** (Feb., 1995) 885–908, arXiv:9406009.

[13] A. F. Heavens and A. N. Taylor, "A spherical harmonic analysis of redshift space," *Mon. Not. R. Astron. Soc.* **275** (July, 1995) 483–497, arXiv:9409027.

[14] K. B. Fisher, C. A. Scharf, and O. Lahav, "A spherical harmonic approach to redshift distortion and a measurement of Omega(0) from the 1.2-Jy IRAS Redshift Survey," *Mon. Not. R. Astron. Soc.* **266** (Jan., 1994) 219, arXiv:9309027.

[15] H. Tadros, W. E. Ballinger, A. N. Taylor, A. F. Heavens, G. Efstathiou, W. Saunders, C. S. Frenk, *et al.*, "Spherical harmonic analysis of the PSCz galaxy catalogue: redshift distortions and the real-space power spectrum," *Mon. Not. R. Astron. Soc.* **305** (May, 1999) 527–546, arXiv:9901351.

[16] A. Rassat and A. Refregier, "3D spherical analysis of baryon acoustic oscillations," *Astron. Astrophys.* **540** (Apr., 2012) A115, arXiv:1112.3100.

[17] C. Shapiro, R. G. Crittenden, and W. J. Percival, "The complementarity of redshift-space distortions and the integrated Sachs-Wolfe effect: a 3D spherical analysis," *Mon. Not. R. Astron. Soc.* **422** (May, 2012) 2341–2350, arXiv:1109.1981 [astro-ph.CO].

[18] B. Leistedt, A. Rassat, A. Réfrégier, and J.-L. Starck, "3DEX: a code for fast spherical Fourier-Bessel decomposition of 3D surveys," *Astron. Astrophys.* **540** (Apr., 2012) A60, arXiv:1111.3591 [astro-ph.CO].

[19] J. Yoo and V. Desjacques, "All-sky analysis of the general relativistic galaxy power spectrum," *Phys. Rev. D* **88** no. 2, (July, 2013) 023502, arXiv:1301.4501.

[20] A. Nicola, A. Refregier, A. Amara, and A. Paranjape, "Three-dimensional spherical analyses of cosmological spectroscopic surveys," *Phys. Rev. D* **90** no. 6, (Sept., 2014) 063515, arXiv:1405.3660.

[21] F. Lanusse, A. Rassat, and J.-L. Starck, "3D galaxy clustering with future wide-field surveys: Advantages of a spherical Fourier-Bessel analysis," *Astron. Astrophys.* **578** (June, 2015) A10, arXiv:1406.5989.

[22] A. Loeb, "The optimal cosmic epoch for precision cosmology," *J. Cosmol. Astropart. Phys.* **5** (May, 2012) 028, arXiv:1203.2622.

[23] D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo, "Observational probes of cosmic acceleration," *Phys. Rep.* **530** (Sept., 2013) 87–255, arXiv:1201.2434.

[24] N. Kaiser, "On the spatial correlations of Abell clusters," *Astrophys. J. Lett.* **284** (Sept., 1984) L9–L12.

[25] N. Kaiser, "Clustering in real space and in redshift space," *Mon. Not. R. Astron. Soc.* **227** (July, 1987) 1–21.

[26] J. Yoo, "Complete Treatment of Galaxy Two-Point Statistics: Gravitational Lensing Effects and Redshift-Space Distortions," *Phys. Rev. D* **79** no. 2, (Jan., 2009) 023517, arXiv:0808.3138.

[27] J. Yoo, A. L. Fitzpatrick, and M. Zaldarriaga, "New perspective on galaxy clustering as a cosmological probe: General relativistic effects," *Phys. Rev. D* **80** no. 8, (Oct., 2009) 083514, arXiv:0907.0707.

[28] J. Yoo, "General relativistic description of the observed galaxy power spectrum: Do we understand what we measure?," *Phys. Rev. D* **82** no. 8, (Oct., 2010) 083508, arXiv:1009.3021.

[29] C. Bonvin and R. Durrer, "What galaxy surveys really measure," *Phys. Rev. D* **84** no. 6, (Sept., 2011) 063505, arXiv:1105.5280 [astro-ph.CO].

[30] A. Challinor and A. Lewis, "Linear power spectrum of observed source number counts," *Phys. Rev. D* **84** no. 4, (Aug., 2011) 043516, arXiv:1105.5292 [astro-ph.CO].

[31] D. Jeong, F. Schmidt, and C. M. Hirata, "Large-scale clustering of galaxies in general relativity," *Phys. Rev. D* **85** (2012) 023504, arXiv:1107.5427.

[32] J. Yoo, "Relativistic effect in galaxy clustering," *Class. Quant. Grav.* **31** no. 23, (Sept., 2014) 234001, arXiv:1409.3223.

[33] J. Yoo, "Proper-time hypersurface of nonrelativistic matter flows: Galaxy bias in general relativity," *Phys. Rev. D* **90** no. 12, (Dec., 2014) 123507, arXiv:1408.5137.

[34] V. Tansella, C. Bonvin, R. Durrer, B. Ghosh, and E. Sellentin, "The full-sky relativistic correlation function and power spectrum of galaxy number counts. Part I: theoretical aspects," *J. Cosmol. Astropart. Phys.* **3** (Mar., 2018) 019, arXiv:1708.00492.

[35] F. Scaccabarozzi, J. Yoo, and S. G. Biern, "Galaxy two-point correlation function in general relativity," *J. Cosmol. Astropart. Phys.* **10** (Oct., 2018) 024, arXiv:1807.09796.

[36] J. Yoo and R. Durrer, "Gauge-transformation properties of cosmological observables and its application to the light-cone average," *J. Cosmol. Astropart. Phys.* **9** (Sept., 2017) 016, arXiv:1705.05839.

[37] C. Bonvin, R. Durrer, and M. A. Gasparini, "Fluctuations of the luminosity distance," *Phys. Rev. D* **73** no. 2, (Jan., 2006) 023523, arXiv:0511183.

[38] C. Bonvin, R. Durrer, and M. Kunz, "Dipole of the Luminosity Distance: A Direct Measure of H(z)," *Physical Review Letters* **96** no. 19, (May, 2006) 191302, astro-ph/0603240.

[39] J. Yoo and F. Scaccabarozzi, "Unified treatment of the luminosity distance in cosmology," *J. Cosmol. Astropart. Phys.* **9** (Sept., 2016) 046, 1606.08453.

[40] S. G. Biern and J. Yoo, "Gauge-invariance and infrared divergences in the luminosity distance," *J. Cosmol. Astropart. Phys.* **4** (Apr., 2017) 045, 1606.01910.

[41] F. Scaccabarozzi and J. Yoo, "Light-cone observables and gauge-invariance in the geodesic light-cone formalism," *J. Cosmol. Astropart. Phys.* **6** (June, 2017) 007, arXiv:1703.08552 [gr-qc].

[42] S. G. Biern and J. Yoo, "Correlation function of the luminosity distances," *J. Cosmol. Astropart. Phys.* (Sept., 2017) 026, 1704.07380.

[43] E. Mitsou, J. Yoo, R. Durrer, F. Scaccabarozzi, and V. Tansella, "Angular n-point spectra and cosmic variance on the light cone," *J. Cosmol. Astropart. Phys.* (May, 2019) , arXiv:1905.01293.

[44] S. Cole, K. B. Fisher, and D. H. Weinberg, "Fourier Analysis of Redshift Space Distortions and the Determination of Omega," *Mon. Not. R. Astron. Soc.* **267** (Apr., 1994) 785, arXiv:astro-ph/9308003.

[45] J. Yoo and U. Seljak, "Wide-angle effects in future galaxy surveys," *Mon. Not. R. Astron. Soc.* **447** (Feb., 2015) 1789–1805, arXiv:1308.1093.

[46] J. C. Mather *et al.*, "Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument," *Astrophys. J.* **420** (Jan., 1994) 439–444.

[47] J. Yoo, N. Grimm, E. Mitsou, A. Amara, and A. Refregier, "Gauge-invariant formalism of cosmological weak lensing," *J. Cosmol. Astropart. Phys.* **4** (Apr., 2018) 029, arXiv:1802.03403.

[48] N. Grimm and J. Yoo, "Jacobi mapping approach for a precise cosmological weak lensing formalism," *J. Cosmol. Astropart. Phys.* **7** (July, 2018) 067, arXiv:1806.00017.

[49] C. Bonvin, "Effect of peculiar motion in weak lensing," *Phys. Rev. D* **78** no. 12, (Dec., 2008) 123530, arXiv:0810.0180.

[50] F. Bernardeau, C. Bonvin, and F. Vernizzi, "Full-sky lensing shear at second order," *Phys. Rev. D* **81** no. 8, (Apr., 2010) 083002, arXiv:0911.2244 [astro-ph.CO].

[51] F. Schmidt and D. Jeong, "Cosmic rulers," *Phys. Rev. D* **86** no. 8, (Oct., 2012) , arXiv:1204.3625.

[52] F. Schmidt and D. Jeong, "Large-scale structure with gravitational waves. II. Shear," *Phys. Rev. D* **86** no. 8, (Oct., 2012) 083513, arXiv:1205.1514.

[53] M. Tegmark, A. N. Taylor, and A. F. Heavens, "Karhunen-Loeve Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?," *Astrophys. J.* **480** (May, 1997) 22, arXiv:9603021.

[54] J. R. Bond, A. H. Jaffe, and L. Knox, "Estimating the power spectrum of the cosmic microwave

background," *Phys. Rev. D* **57** (Feb., 1998) 2117–2137, arXiv:9708203.

[55] A. Heavens, "3D weak lensing," *Mon. Not. R. Astron. Soc.* **343** (Aug., 2003) 1327–1334, arXiv:0304151.

[56] M. Tegmark *et al.*, "Cosmological constraints from the SDSS luminous red galaxies," *Phys. Rev. D* **74** no. 12, (Dec., 2006) 123507, arXiv:astro-ph/0608632.

[57] H.-J. Seo and D. J. Eisenstein, "Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys," *Astrophys. J.* **598** (Dec., 2003) 720–740, arXiv:astro-ph/0307460.

[58] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, *et al.*, "Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies," *Astrophys. J.* **633** (Nov., 2005) 560–574, arXiv:astro-ph/0501171.

[59] I. Ben-Dayan, M. Gasperini, G. Marozzi, F. Nugier, and G. Veneziano, "Average and dispersion of the luminosity-redshift relation in the concordance model," *J. Cosmol. Astropart. Phys.* **6** (June, 2013) 2, arXiv:1302.0740.

[60] I. Ben-Dayan, R. Durrer, G. Marozzi, and D. J. Schwarz, "Value of H_0 in the Inhomogeneous Universe," *Phys. Rev. Let.* **112** no. 22, (June, 2014) 221301, arXiv:1401.7973.

[61] L. Hui and P. B. Greene, "Correlated fluctuations in luminosity distance and the importance of peculiar motion in supernova surveys," *Phys. Rev. D* **73** no. 12, (June, 2006) 123526, astro-ph/0512159.

[62] J. Yoo and others, "Maximum cosmological information contents in supernova observations (in preparation)," (2019) .

[63] J. Yoo, E. Mitsou, Y. Dirian, R. Durrer, and M. Zaldarriaga, "$\bar{t}$: a new cosmological parameter?" *in preparation* (2019) .

[64] Y. Mellier, "Probing the Universe with Weak Lensing," *Ann. Rev. Astron. Astrophys.* **37** (1999) 127–189, arXiv:9812172.

[65] M. Bartelmann and P. Schneider, "Weak gravitational lensing," *Phys. Rep.* **340** (Jan., 2001) 291–472, arXiv:9912508.

[66] A. Refregier, "Weak Gravitational Lensing by Large-Scale Structure," *Ann. Rev. Astron. Astrophys.* **41** (2003) 645–668, arXiv:0307212.

[67] D. Munshi, P. Valageas, L. van Waerbeke, and A. Heavens, "Cosmology with weak lensing surveys," *Phys. Rep.* **462** (June, 2008) 67–121, astro-ph/0612667.

[68] The Dark Energy Survey Collaboration, "The Dark Energy Survey," arXiv:astro-ph/0510346.

[69] J. T. A. de Jong *et al.*, "The first and second data releases of the Kilo-Degree Survey," *Astron. Astrophys.* **582** (Oct., 2015) A62, arXiv:1507.00742.

[70] A. Heavens, A. Refregier, and C. Heymans, "Intrinsic correlation of galaxy shapes: implications for weak lensing measurements," *Mon. Not. R. Astron. Soc.* **319** (Dec., 2000) 649–656, astro-ph/0005269.

[71] C. M. Hirata, R. Mandelbaum, U. Seljak, J. Guzik, N. Padmanabhan, C. Blake, J. Brinkmann, T. Budávari, A. Connolly, I. Csabai, R. Scranton, and A. S. Szalay, "Galaxy-galaxy weak lensing in the Sloan Digital Sky Survey: intrinsic alignments and shear calibration errors," *Mon. Not. R. Astron. Soc.* **353** (Sept., 2004) 529–549, arXiv:0403255.

[72] W. Hu, "Power Spectrum Tomography with Weak Lensing," *Astrophys. J. Lett.* **522** (Sept., 1999) L21–L24, astro-ph/9904153.

[73] N. Grimm and others, "Cosmological information contents in weak lensing (in preparation)," (2019) .

[74] M. Fukugita, C. J. Hogan, and P. J. E. Peebles, "The Cosmic Baryon Budget," *Astrophys. J.* **503** (Aug., 1998) 518–530, astro-ph/9712020.

[75] J. M. Shull, B. D. Smith, and C. W. Danforth, "The Baryon Census in a Multiphase Intergalactic Medium: 30% of the Baryons May Still be Missing," *Astrophys. J.* **759** (Nov., 2012) 23, arXiv:1112.2706.

[76] F. Nicastro, S. Mathur, M. Elvis, and J. a. Drake, "The mass of the missing baryons in the X-ray forest of the warm-hot intergalactic medium," *Nature* **433** (Feb., 2005) 495–498.

[77] F. Nicastro, J. Kaastra, Y. Krongold, S. Borgani, *et al.*, "Observations of the missing baryons in the warm-hot intergalactic medium," *Nature* **558** (June, 2018) 406–409, arXiv:1806.08395.

[78] J. Yoo and others, "Cosmic variance contribution to the missing baryon problem (in preparation)," (2019) .

[79] D. N. Limber, "The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. II.," *Astrophys. J.* **119** (May, 1954) 655.

[80] M. LoVerde and N. Afshordi, "Extended Limber Approximation," *ArXiv e-prints* (Sept., 2008) , 0809.5112.

[81] J. Yoo and others, "Galaxy power spectrum analysis on the light cone (in preparation)," (2019) .

[82] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, "The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set," *Astrophys. J.* **473** (Dec., 1996) 576, astro-ph/9605054.

[83] D. J. Fixsen, "The Temperature of the Cosmic Microwave Background," *Astrophys. J.* **707** (Dec., 2009) 916–920, arXiv:0911.1955.

[84] Planck Collaboration *et al.*, "Planck 2018 results. VI. Cosmological parameters," *arXiv e-prints* (July, 2018) , arXiv:1807.06209.

[85] U. Seljak, "Extracting Primordial Non-Gaussianity without Cosmic Variance," *Phys. Rev. Let.* **102** no. 2, (Jan., 2009) 021302, arXiv:0807.1770.

[86] P. McDonald and U. Seljak, "How to evade the sample variance limit on measurements of redshift-space distortions," *J. Cosmol. Astropart. Phys.* **10** (Oct., 2009) 7, arXiv:0810.0323.